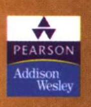
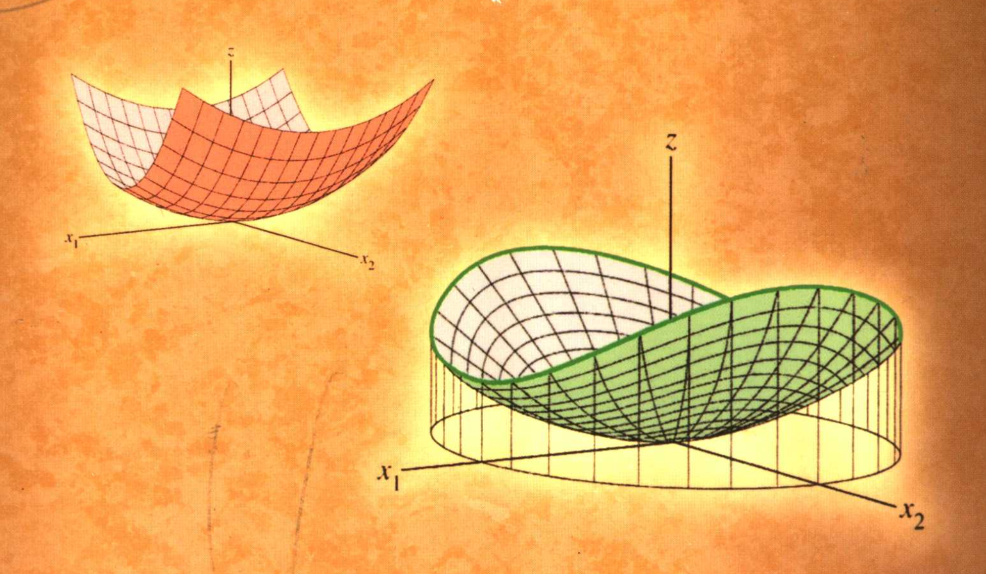
图灵数学·统计学丛书 12





Linear Algebra and Its Applications

线性代数及其应用

（第3版修订版）

［美］David C. Lay 著

沈复兴 傅莺莺 莫单玉 等 译

人民邮电出版社

POSTS & TELECOM PRESS

线性代数及其应用（第3版修订版）

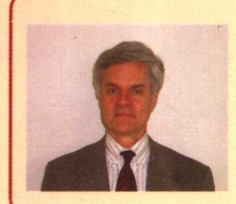
Linear Algebra and Its Applications

“我们相信，本书的出现能够让更多的教师和学生在轻松愉悦的教与学中把握线性代数课程改革的世界潮流、分享线性代数的乐趣，这必将有力地推动我国线性代数课程的改革。”

——沈复兴（北京师范大学）

线性代数作为大学数学的核心课程，其重要性不言而喻。但是多年来，传统的教学方式往往局限于理论体系本身，严密性有余，而启发性和趣味性不足，对线性代数的方法和应用尤其是数值计算缺乏应有的重视。

本书作为美国国家级“线性代数课程研究组”研究成果的体现，是目前全球高等院校采用率最高的优秀线性代数教材之一，全面反映了当前国际教学改革的潮流。先进的教学思想，循序渐进的章节安排，丰富的例题、案例和教学辅助手段，完整的理论推导和证明，以及与各学科领域应用的紧密结合，使本书广受赞誉。本版在第3版的基础上又增加了许多新的内容，包括内容丰富的配套CD，为课程教学提供更多支持，适用于各专业的教学。



David C. Lay 国际知名的杰出数学教育家和学者，马里兰大学教授。他是美国国家科学基金会资助的“线性代数课程研究组”创始人之一，也是美国线性代数课程改革的领导者之一，曾经获得美国数学协会授予的杰出数学教学奖。除本书外，他还与人合著了泛函分析、微积分等方面的教材，也深受好评。

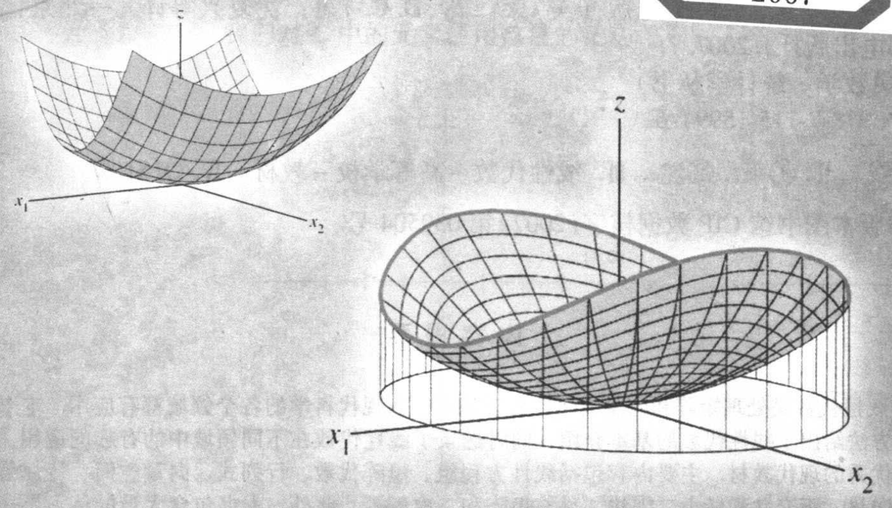
www.PearsonEd.com

本书相关信息请访问：
图灵网站 http://www.turingbook.com
读者/作者热线：(010) 88593802
反馈/投稿/推荐信箱：contact@turingbook.com

分类建议 数学/基础数学

人民邮电出版社网址 www.ptpress.com.cn

ISBN 978-7-115-15994-6/O1
定价：59.00 元 （附光盘）



Linear Algebra and Its Applications

线性代数及其应用

（第3版修订版）

[美] David C. Lay 著

傅莺莺 莫单玉 等 译

人民邮电出版社
北 京

图书在版编目(CIP)数据

线性代数及其应用：第 3 版/(美)莱(Lay,D. C.)著；沈复兴等译. —北京：人民邮电出版社，2007. 7
(图灵数学·统计学丛书)
ISBN 978-7-115-15994-6

Ⅰ. 线… Ⅱ. ①莱…②沈… Ⅲ. 线性代数－高等学校－教材 Ⅳ. O151. 2

中国版本图书馆 CIP 数据核字(2007)第 039504 号

内 容 提 要

线性代数是处理矩阵和向量空间的数学分支，在现代科学的各个领域都有应用. 本书用现代方法给出了线性代数的基本介绍，同时选录了线性代数在不同领域中的有趣的应用，是一本优秀的现代教材. 主要内容包括线性方程组、矩阵代数、行列式、向量空间、特征值与特征向量、正交性和最小二乘法、对称矩阵和二次型等. 此外，本书包含大量的练习题、习题、例题等，便于读者学习、参考.

本书适合作为高等院校理工科相关专业线性代数课程的教材，也可作为相关研究人员的参考书.

图灵数学·统计学丛书

线性代数及其应用（第 3 版修订版）

◆ 著　　[美] David C. Lay
　译　　沈复兴　傅莺莺　莫单玉　等
　责任编辑　明永玲　于越华

◆ 人民邮电出版社出版发行　北京市崇文区夕照寺街 14 号
　邮编　100061　电子函件　315@ptpress.com.cn
　网址　http://www.ptpress.com.cn
　北京铭成印刷有限公司印刷
　新华书店总店北京发行所经销

◆ 开本：700 × 1000　1/16
　印张：28.75
　字数：682 千字　　2007 年 7 月第 1 版
　印数：1－4 000 册　　2007 年 7 月北京第 1 次印刷

著作权合同登记号　图字：01-2006-0315 号

ISBN 978-7-115-15994-6/O1

定价：59.00 元（附光盘）

读者服务热线：(010)88593802　印装质量热线：(010)67129223

版 权 声 明

译者介绍

沈复兴 江苏常熟人，北京师范大学教授，博士生导师．热爱教学，长期从事数理逻辑和计算机科学等领域的研究与教学工作．曾获1986年国家教委科技进步一等奖、1996年国家教委科技进步三等奖、北京师范大学2002年“十佳教师”、2003年“本科生教学名师奖”、2004年宝钢优秀教师奖、2004年北京市教学成果（高教）一等奖、国家级教学成果二等奖、2005年北京师范大学首届钱媛教育基金优秀教师奖．曾任北京师范大学教务长、信息科学学院院长，教育部教学指导委员会计算机科学与技术教学指导委员会非计算机专业计算机基础教学指导分委员会委员．现任中国数学会数理逻辑分会副理事长、秘书长，全国高等师范学校计算机教育研究会荣誉副理事长，全国高等院校计算机基础教育研究会师范专业委员会主任，美国 *Mathematical Reviews* 评论员．

傅莺莺 北京师范大学数学科学学院博士研究生，师从沈复兴教授，研究方向为数理逻辑，主要研究数理逻辑在代数、计算机等领域的应用，发表学术论文近十篇．长期担任北京师范大学公共数学、高等代数、抽象代数、数理逻辑等本科、硕士课程的助教，曾经参与云南省高中信息技术教材、全国软件水平考试辅导材料的编纂，对教材、教法有较深刻的认识．

译 者 序

线性代数是研究矩阵和向量空间的一个数学分支，也是高等学校理、工、经、管类各专业的一门基础课程，在数学、力学、物理学、经济学和其他学科中都有重要应用. 但是，当前线性代数课程的教学效果不尽如人意. 究其原因，主要在于多数教材仍然停留在传统模式，未能适应新的社会需求. 传统的线性代数教材过分追求逻辑的严密性和理论体系的完整性，重理论而轻实践，剥离了概念、原理和范例的几何背景与现实意义，导致教学内容过于抽象，也不利于与其他课程及学生自身专业的衔接，进而造成学生“学不会，用不了”的尴尬局面. 因此，要改革线性代数课程教学，首要任务就是开发或者引进一批优秀的现代线性代数教材.

David C. Lay 教授是美国线性代数课程研究项目的创始人之一，一直致力于倡导数学课程的现代化. 他从事线性代数的科研与教学工作多年，迄今已发表学术论文 30 余篇，与他人合著多本数学教材，并且四次荣获大学教学优秀奖. David C. Lay 教授编撰的这本教材自第 1 版起就得到了广泛的关注和认可，其第 3 版的影印版和中文译本也相继在国内发行和使用，效果良好. 此次由人民邮电出版社引进了该书第 3 版修订版，它在第 3 版的基础上又增加了许多新的内容，包括配套 CD 等. 与传统教材相比，本教材风格清新活泼，其主要特点有：

注重几何背景，从具体看抽象 要讲清过于抽象的概念是线性代数教学中的主要难点. 本教材强调线性代数的实际背景，将其与空间解析几何建立起联系，抓住“向量”这一基本概念，将线性代数视为空间解析几何在 n 维空间的推广.

紧密联系实际，服务专业课程 本教材包含丰富的应用范例，取材广泛. 这些实际问题(如经济投入产出问题、飞机制造模型、猫头鹰种群问题)既为学生理解概念提供了认识基础，也有助于加强与专业课程的联系，使学生学有所用.

运用现代技术，培养实际能力 本教材着重培养学生运用线性代数理论解决实际问题的数学建模能力，要求他们理解和掌握基本算法，但不提倡将过多精力花在计算上. 配套的 CD 中包括常用数学软件、工具的使用方法，也提供了教材中全部数值计算的数据，大大节省了计算的时间，使学生能够专注于理解概念和算法本身.

课程体系完善，配套资源丰富 配套 CD 和网页提供了详尽的学习指南，指导学生如何学习这门课程. CD 和网页中的软件工具和数据也完全可以用来开设配套的实验课程.

此外，我们在译文中添加了若干“译者注”，以帮助读者抛开文化背景的差异，更好地理解原文. 我们相信，修订版中译本的出现能够让更多的教师和学生在轻松愉悦的教、学中把握线性代数课程改革的世界潮流，分享线性代数的乐趣，这必将有力地推动我国线性代数课程的改革.

参与本书翻译及校对的还有北京师范大学数学科学学院、信息科学学院的研究生隋建宝、马慧芳、王慎玲、李永强和马鑫. 翻译过程中得到了人民邮电出版社图

灵公司数学编辑的很大帮助，在此一并表示感谢. 由于译者水平有限，书中难免疏漏和不妥之处，敬请广大师生、同行专家批评指正！

译 者

北京师范大学数学科学学院

前　　言

广大师生对本书前三版的评价很高，第 3 版修订版在此基础上为课程教学以及软件技术应用提供了更多支持．与前面几版一样，本书用现代方法给出了线性代数的基本介绍，同时选录了线性代数在不同领域中的有趣的应用．本书适用于已顺利完成两学期大学数学课程（通常是微积分）的学生．

本书的目标是帮助学生掌握基本的概念和技能，这些知识在他们以后的工作中还将用到．本书的选题采纳了线性代数课程研究小组的建议，建立在对学生实际需求的细致调查基础上，并且得到相关学科专业人士的一致认可．希望本课程能够成为对本科生最有用且最有趣的数学课程之一．

特　　色

关键概念的及早引入

本书前七章以 $\mathbf{R}^n$ 为背景介绍了线性代数的许多基本思想，后续章节中则逐步考察它们在不同背景下的实质．基本概念的概括和推广都是这些思想的自然延伸，借助第 1 章所建立的几何直观，我们可以将其形象化以加深理解．我认为，本书最大的优点在于整体难度适当．

矩阵乘法的现代观点

一套好的记号非常重要，本书采用的是科学家和工程师们运用线性代数时所用的记号．书中的定义及证明都基于矩阵的列而非元素展开，其核心思想是将矩阵－向量乘积 $\boldsymbol{Ax}$ 视为 $\boldsymbol{A}$ 中列的一个线性组合．这种现代观点简化了许多论证，同时也为向量空间与线性方程组建立了联系．

线性变换

线性变换是贯穿全书始末的一个脉络，这一概念的引入给本书抹上了浓重的几何色彩．例如，在第 1 章中，线性变换为我们理解矩阵向量乘法提供了一个动态的、图形化的视角．

本征值和动态系统

本征值出现在本书第 5 章和第 7 章．这部分教材内容的教学需要几周的时间，因此学生有充足的时间来消化和复习这些重要概念．本征值源于并且应用于离散或连续的动态系统，这类动态系统在本书 1.10 节、4.8 节、4.9 节以及第 5 章的五节中都有介绍．实际教学安排中，如果不选用第 4 章，而代之以 2.8 和 2.9 两节，那么五周后就可以讲到第 5 章．可选的这两节从第 4 章中提取了第 5 章要用到的关于向量空间的全部概念．

正交性与最小二乘问题

与线性代数其他入门教材相比，本书对正交性与最小二乘问题的介绍要全面、系统得多．线性代数课程研究小组强调有必要将这部分内容组织成一个完整单元，这不仅仅因为正交性在计算机计算和数值线性代数中有着十分重要的地位，而且也

考虑到实际工作中常常需要处理不相容的线性方程组.

教学特色

应用

本书从工程学、计算机科学、数学、物理学、生物学、经济学和统计学等众多学科中广泛选用了一些实际应用的例子，揭示了线性代数在其中解释基本原理、简化计算等方面起到的重要作用. 部分应用实例独立成节，其余则以例题或习题的形式给出. 本书每一章开头的实例介绍都为线性代数的一类应用做了铺垫，有助于激发学生学习后继数学知识的兴趣. 在一章临近结束的时候，又重新回到实例介绍中的问题并加以解决.

突出几何重点

几何直观的建立常常有助于学生理解抽象的概念，因此本书对每一个重要概念都配以几何解释. 本书的插图比一般的线性代数教材更为丰富，而且其中一部分在其他教材中从未出现过.

例题

本书提供了许多例题，并且配以详尽的说明材料，这在很多线性代数教材中是没有的. 教师按通常的课程计划一般无法讲完全部例题，不过由于例题写得清楚、详细，学生完全可以自学.

定理及证明

本书的重要结果都以定理形式给出. 为方便查阅，在有灰色底纹的方框中还列出了其他一些常用的事实. 许多定理都有正式的证明，初学者都能看懂；另有一些定理，其证明的主要过程在例题中给出；其余定理证明较为简单，留给学生作为练习，我们认为这样安排对他们十分有益.

基础练习

每个习题集之前都精心准备了若干基础练习，并且在习题集后给出其详细解答. 这些练习主要针对习题集中潜在的难点而设置，相当于做习题前的“热身”，其解答往往为习题提供了有用的提示或警告.

习题

本书包含了丰富的习题，有简单的计算，也有需要略加思考的概念题. 从事教学多年，我从学生的作业、试卷中总结了一些较难理解的概念. 针对这些难点，书中专门设置了大批原创性问题. 习题的编排顺序与教材顺序一致，每讲完一部分内容就可以布置家庭作业. 习题的一个显著特点是数值简单、问题直接明了，所以学生不必在计算上浪费时间. 总体来说，习题着重于理解而非机械计算.

判断题

为了鼓励学生阅读全书并且培养他们缜密思考的习惯，我在全书 33 节中设置了 300 道简单判断题，数量仅次于计算题. 这些习题的答案可从教材中直接查到，它们为学生完成后面的概念题做了铺垫. 学生们认真阅读教材后将会十分喜爱这一类习题. 基于课堂检验的效果以及与学生的广泛交流，我决定不把习题答案收录于书中. (学习指南给出了奇数号习题的解答) 另外还有 150 道判断题(大部分在每一章末尾)

用于检验学生对教材内容的理解程度. 对于这些习题，教材中大都给出了简单的答案，但没有给出理由(通常需要学生进一步思考).

写作题

书写连贯的数学证明，不仅是对准备继续深造的学生，而且是对学习线性代数的全体学生的一个基本要求. 本书中许多习题的答案都要求包括书面说明. 概念题通常要求给出简短的证明，且常常伴以解题提示. 对于所有奇数编号的习题，其答案或者在本书最后给出，或者书后仅给出提示. 详细解答另见*学习指南*.

关于数值计算

本书十分强调计算机对线性代数在科学、工程学中的发展和应用所产生的影响. 书中多次出现的注记讨论了数值计算中需要注意的一些问题，以及理论概念(例如矩阵求逆)与计算机实现(例如 LU 分解)之间的差异.

配套 CD 及网页支持

学习指南

本书修订版的配套 CD 中包含完整的*学习指南*. *学习指南*也是这门课程的一个主要部分. 教材中的图标(SG)提示学生，为了理解教材中的关键概念可以参考*学习指南*的哪几节. 奇数号习题中，每隔两道题的第三题有详细解答，可以在*学习指南*中找到，用以帮助学生检查作业，*学习指南*还为那些书后只包含提示的奇数号写作题提供了完整答案. 多次出现的“警告”用于提醒学生易犯的错误以及如何避免出错. 标记 MATLAB 的方框内介绍了一些需要使用的命令. *学习指南*的附录内包含 Maple、Mathematica 以及 TI 和 HP 图形计算器的对照信息.

软件的使用

如果课程中需要使用 Maple、Mathematica 或者 TI 和 HP 图形计算器，可以参考 CD 中的软件使用介绍(见第 1 章最后一页). 此外，*学习指南*中也为第一次接触上述软件的用户提供了详细说明.

数据文件

CD 中包含数百个数据文件，覆盖了教材中约 900 道数值计算题以及案例研究、应用项目的全部数据. 数据文件以多种格式给出，分别针对 MATLAB、Maple、Mathematica、TI－83＋/86/89 和 HP48G 图形计算器. 对任一问题，学生们只需通过少量的键盘操作就可以得到矩阵和向量，从而避免了数据录入错误，也节约了时间.

新的 MATLAB 项目

这些探究性项目可以引导学生发现线性代数中一些基本的数学和数值问题. Rick Smith 最初开发这些项目的目的，是用于辅助佛罗里达大学计算线性代数这门课程的教学，他们使用《线性代数及其应用》这本教材已有多年. 教材中使用图标(CD)引用这类项目. 探究性项目中，约有一半是关于基本概念的，例如列空间、对角化和正交投影；另一些讨论数值计算问题，比如浮点运算、迭代法和 SVD；还有一些涉及应用，例如拉格朗日插值和马尔可夫链.

www. laylinalgebra. com

本书的网站提供了配套 CD 中除*学习指南*及新的 MATLAB 项目以外的全部材料.

此外，网站上还包括修订版教材的第1章以及学习指南的第1章，其目的是在一开学教材脱销时保障老师正常开课. 网站为学生提供了**复习表**和**模拟试卷**(附答案)，其范围覆盖教材中的主要内容，并且都是我从多年的教学工作中总结提炼出来的. 每个复习表都详细列出了关键的定义、定理以及教材中使用的技巧.

基于各章的应用

网站包括7个案例研究，都是每章开头实例介绍的拓展，其中添加了一些现实生活中的数据以期进行更深入的研究. 此外，网站上还列出了20多个应用项目，其中一部分是教材内容的补充，另一部分介绍了新的应用，例如三次样条、飞机飞行路线、体育竞赛中的占优矩阵以及纠错码. 其中有些数学应用需要综合运用多项式根的位置、圆锥截面、二次曲面、二元函数的极值等知识，另外也涉及数值线性代数的知识，例如条件数、矩阵分解以及求本征值的QR方法. 以上实例研究及应用项目均配有习题，其所涉及的数据集较大(因而需要借助软件来求解).

致 谢

我衷心感谢多年以来在各方面给予我帮助，使本书最终得以付梓的人.

感谢Israel Gohberg和Robert Ellis与我共同研究线性代数长达15年之久，他们的思想深深影响了我对线性代数的看法.

我也很荣幸能和David Carlson、Charles Johnson以及Duane Porter一同在线性代数课程研究小组工作，他们关于线性代数教学的思想在本书许多地方都有体现.

我还要诚挚地感谢以下审阅者，感谢他们的细致分析和所提的建设性意见:

第3版的审阅者及课堂检验者

大河谷州立大学的David Austin，德州大学奥斯汀分校的G. Barbanson，Valdosta州立大学的Ashok Kumar，康奈尔大学的Kenneth Brown，加州州立大学萨克拉门托分校的Earl Kymala，圣迭戈州立大学的David Carlson，明尼苏达大学杜鲁斯分校的Kathryn Lenz，杨百翰大学的Greg Conner，赛拉丘兹大学的Jaques Lewin，马里兰大学的Casey T. Cremins，托莱多大学的En-Bing Lin，Okanagan大学学院的Sylvie DesJardins，马里兰大学的Andrei Maltsev，南亚拉巴马大学的Daniel Flath，拿骚公共学院的Abraham Mantell，俄亥俄州立大学的Yuval Flicker，马里兰大学巴尔的摩县分校的Madhu Nayakkankuppam，克拉克森大学的Scott Fulton，斯坦福大学的Lei Ni，Drexel大学的Herman Gollwitzer，宾州州立大学的Gleb Novitchkov，科罗拉多大学科罗拉多泉分校的Jeremy Haefner，南佛罗里达大学的Ralph Oberste-Vorth，佛罗里达大学的William Hager，布朗大学的Dev Sinha，北亚利桑那大学的John Hagood，圣何塞州立大学的Wasin So，科罗拉多矿校的Willy Hereman，俄亥俄州立大学的Ron Solomon，科罗拉多州立大学的Alexander Hulpke，康涅狄格大学的Eugene Spiegel和Alan Stein，惠特曼学院的Doug Hundley，科罗拉多州立大学的James Thomas，康涅狄格大学的James F. Hurley，Bethel学院的Brian Turnquist，路易斯安那州立大学的Jurgen Hurrelbrink，西俄勒冈大学的Michael Ward，拉瓜迪亚公共学院的Jerry G. Ianni，亚利桑那州立大学的Bruno Welfert，德州大学奥斯汀分校的Jack Xin.

对于本次修订，我要感谢Winthrop大学的Thomas Polaski帮我修改了第3版的补

充材料，并且不断提供一些好的建议. 我也要感谢佛罗里达大学的 Rick Smith 同意将他设计的 MATLAB 项目收录于修订版中，同时感谢 Taylor 大学的 Jeremy Case 帮助审阅和修改这些项目. 最后，我要感谢 Addison-Wesley 的全体工作人员为修订版所付出的辛勤劳动.

David C. Lay

致 学 生

在你即将学习的所有大学本科数学课程中，这门课程很可能是最有趣且最值得努力的课程．事实上，有的学生毕业后在信中或者当面告诉我，他们在一些大公司或工程研究所工作时，有时仍然会使用这本教材作为参考．以下列出几条实用的建议和信息，希望能够帮助你掌握这门课程的内容并且享受其中的乐趣．

在线性代数中，概念和计算同等重要．每个习题集的开头都安排了简单的数值计算题，目的是帮助你检查对基本算法的理解．在今后的工作中，你当然可以使用计算机来完成这些计算，但你必须学会选择算法，理解计算结果的意义并且向他人解释清楚．正因如此，这本教材的许多习题都要求你解释或者验证计算过程，完整的习题答案必须包含书面的解释．对于奇数号的习题，你可以在答案部分找到合理的解释，或者至少是一个好的提示．不过，你应先尝试独立完成习题，尽量不看答案，便于发现哪些知识自己还没有真正理解．

只有反复、认真地阅读教材，你才能真正掌握线性代数的概念．新出现的术语以黑体字标识，有时也会在方框内予以定义．术语表在教材的末尾．为便于参考，重要的事实一般以定理的形式陈述，或者置于有灰色底纹的方框内．我建议你读完前言的前四页，了解本书的结构，这将有助于你理解课程是如何展开的．

从某种意义上说，线性代数就是一门语言．你必须像对待外语一样，每天都学习线性代数．为了真正理解教材中某一节的内容，你往往需要完全掌握前面章节的内容和习题．跟上课程的进度可以节省很多时间并且避免很多麻烦．

注记

即便你没有使用计算机或图形计算器，我也希望你能阅读教材中的注记．在现实生活中，线性代数的大多数应用都涉及数值计算，其中或多或少都包含一些数值误差，尽管这个误差可能非常小．书中的注记提醒你在今后应用线性代数时需要注意的问题，如果你现在学习了注记，今后就更容易记起它们．

如果你喜欢阅读注记，或许会愿意继续学习数值线性代数这门课程．为了满足对更强大的计算能力的迫切需求，计算机科学家和数学家都在研究数值线性代数，希望找到更高速、更可靠的计算算法，同时，电机工程师也在设计更高速、更小巧的计算机以实现这些算法．这是一个令人激动而振奋的研究领域，你的第一门线性代数课程将帮助你迎接这一挑战．

学习指南

为了帮助你学好这门课程，教材的配套 CD 中提供了完整的*学习指南*．它不但能帮助你学习线性代数，而且也将指导你如何学习数学．教材关键部分所出现的图标 (SG) 提示你查阅*学习指南*中标题为“掌握线性代数的概念”的小节，这些小节针对怎样建立关键概念的复习表提供了很好的建议．准备这样的复习表是学好线性代数的秘诀之一，因为这有助于你建立起*知识之间的联系*．这些联系将为你建立牢固的

基础，使你能够掌握并且牢记课程中的主要概念.

奇数号习题中，每隔两道题的第三题有详细解答，可以在学习指南中找到，此外学习指南还为只在书后答案部分给出提示的奇数号写作题提供了完整答案. 将指南与教材分开，是因为你必须学会独立完成习题(根据我多年的经验，对于大多数学生来说，经常查阅教材后的答案对学习数学有负面影响). 指南中也包含了对常见错误的警告，对重要习题或者考点的有益提示.

如果你会使用 MATLAB、Maple、Mathematica，或者 TI、HP 图形计算器，那么你做作业时将节省很多时间. 学习指南就是你的“实验手册”，它会告诉你如何利用这些矩阵软件或工具，其中包括一些新命令的介绍. 你可以在配套 CD 中下载教材中 850 多道习题的数据(只需少量的键盘输入，你就可以在屏幕上完成任何一道数值作业题). 专用的矩阵命令将代你完成计算.

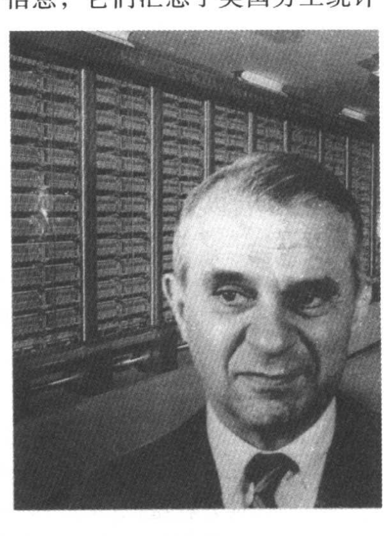
良好的开端是成功的一半. 你在这门课程最初几周内所做的工作，将决定你整学期的学习进度以及最后的掌握程度. 请尽快阅读学习指南中的“如何学习线性代数”. 我的学生们发现其中的学习方法很有用，希望对你也一样.

目　录

第 1 章　线性方程组

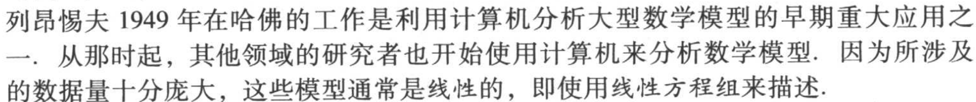
实例介绍：经济学和工程学中的线性模型

1949 年夏末，哈佛大学的瓦·列昂惕夫教授小心翼翼地将最后一张穿孔卡片插入学校的 Mark Ⅱ计算机. 这些卡片存储着美国经济信息，它们汇总了美国劳工统计署历时两年紧张工作所得的250 000 多条数据. 列昂惕夫把美国的经济系统分成500 个“部门”，如煤炭工业、汽车工业、通讯业等. 针对每个部门给出了一个线性方程，描述该部门如何向其他部门分配产出. 但是，由于 Mark Ⅱ(当时最大的计算机之一)还不能处理含 500 个未知量、500 个方程的方程组，所以列昂惕夫把这个问题提炼成一个只含 42 个未知量、42 个方程的方程组.



在 Mark Ⅱ上编程求解列昂惕夫的这 42 个方程历时数月，他急于知道计算机解决这个问题要花费多长时间. 最后，经过 56 个小时的持续运转，Mark Ⅱ终于求出了一个解. 我们将在 1.6 和 2.6 节中讨论这个解的性质.

列昂惕夫开启了通往经济学数学模型一个新时代的大门，并于 1973 年荣获诺贝尔经济学奖. 列昂惕夫 1949 年在哈佛的工作是利用计算机分析大型数学模型的早期重大应用之一. 从那时起，其他领域的研究者也开始使用计算机来分析数学模型. 因为所涉及的数据量十分庞大，这些模型通常是线性的，即使用*线性方程组*来描述.



线性代数在应用上的重要性与计算机的计算性能成正比例增长，而这一性能伴随着计算机每一代硬、软件的革新在不断提升. 最终，计算机并行处理和大规模计算的迅猛发展将计算机科学与线性代数紧紧地联系在一起.

科学家和工程师如今处理的问题远比几十年前想象的要复杂得多. 今天，对于理工和商业专业的学生来说，线性代数比其他大学本科数学课程具有更大的潜在价值. 本书的内容，将为你深入研究许多有趣的领域奠定基础. 这里列举一些可能的领域，其他的将在后面介绍.

- *石油勘探*. 当船只勘查海底石油储量时，船上的计算机每天都要计算数千个线性方程组. 方程组的震动数据从气枪发射所产生的水下冲击波中获取. 冲击波经海底岩石反射，被连接在船尾数英里外的地震探波仪接收并测量.
- *线性规划*. 当今，很多重要的管理决策建立在含有上百个变量的线性规划模型上. 例如，航空工业利用线性规划方法制定机组人员日程表，监控飞机飞行位置，或者制定各式各样的支撑服务计划，如维护、登机口作业等.

- 电网．工程师利用仿真软件设计电路以及包含数百万晶体管的微芯片．这类软件离不开线性代数方法和线性方程组．

线性方程组是线性代数的核心，本章将借助线性方程组简单而具体地介绍线性代数中的核心概念．1.1和1.2节介绍求解线性方程组的一套系统方法．这个计算方法将贯穿全书．1.3和1.4节说明线性方程组如何等价于向量方程和矩阵方程．这种等价性将向量线性组合问题化简为线性方程组问题．本章后半部分学习的张成、线性无关和线性变换等基本概念，将在全书发挥重要作用，深入理解它们将有助于我们感受线性代数的力与美．

1.1 线性方程组

一个包含变量 $x_1, x_2, \cdots, x_n$ 的**线性方程**(linear equation)是具有如下形式的方程：

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{1}$$

其中 b 与**系数**(coefficients) $a_1, a_2, \cdots, a_n$ 均为实数或复数，它们通常是已知的．下标 n 是任意的正整数．在本书的例题和习题中，n 通常是2到5之间的数．而在实际问题中，n 可以是50或5000，甚至更大．

2

方程

$$4x_1 - 5x_2 + 2 = x_1 \text{和} x_2 = 2(\sqrt{6} - x_1) + x_3$$

都是线性的，因为它们重新整理后可以得到方程(1)的形式：

$$3x_1 - 5x_2 = -2 \text{ 和 } 2x_1 + x_2 - x_3 = 2\sqrt{6}$$

方程

$$4x_1 - 5x_2 = x_1x_2 \text{和} x_2 = 2\sqrt{x_1} - 6$$

不是线性方程，因为这两个方程中分别出现了 x_1x_2 和 $\sqrt{x_1}$．

一个**线性方程组**(system of linear equations)或**线性系统**(linear system)是一个或多个线性方程的集合，其中每个方程都含有相同的变量，比如说，$x_1, x_2, \cdots, x_n$．例如

$$\begin{aligned} 2x_1 - x_2 + 1.5x_3 &= 8 \\ x_1 \qquad - 4x_3 &= -7 \end{aligned} \tag{2}$$

方程组的**解**(solution)是一列数 $(s_1, s_2, \cdots, s_n)$，使得用 $s_1, \cdots, s_n$ 分别替换方程中的 $x_1, \cdots, x_n$ 后，方程组中的每个方程都成立．例如，$(5, 6.5, 3)$ 是方程组(2)的一个解，因为用这些值分别替换(2)中的 x_1，x_2 和 x_3，方程化简成 $8 = 8$ 和 $-7 = -7$．

线性方程组所有可能的解构成的集合称为它的**解集**(solution set)．如果两个线性方程组有相同的解集，则称这两个方程组**等价**(equivalent)，即第一个方程组的每个解都是第二个方程组的解，第二个方程组的每个解都是第一个方程组的解．

对于包含两个变量两个方程的线性方程组来说，其解集较容易求出，因为这相当于求两条直线的交点．一个典型的问题为：

$$\begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 3x_2 &= 3 \end{aligned}$$

这两个方程的图像是两条直线，记为 l_1，l_2．数对 (x_1, x_2) 同时满足两个方程当且仅当点 (x_1, x_2) 同时落在 l_1，l_2 这两条直线上．容易验证，上述方程组的解是单点 $(3, 2)$．见图1-1．

3

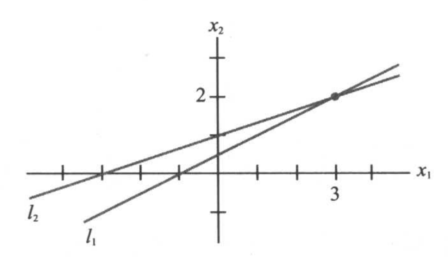


图 1-1 唯一解

当然，两条直线不一定相交于一点——它们可以平行，即没有交点，或者重合，即相交于直线上的每一点. 图 1-2 给出了这些图像，它们对应于下列方程组：

$$\text{(a)}\quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= 3 \end{aligned} \qquad \text{(b)}\quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= 1 \end{aligned}$$

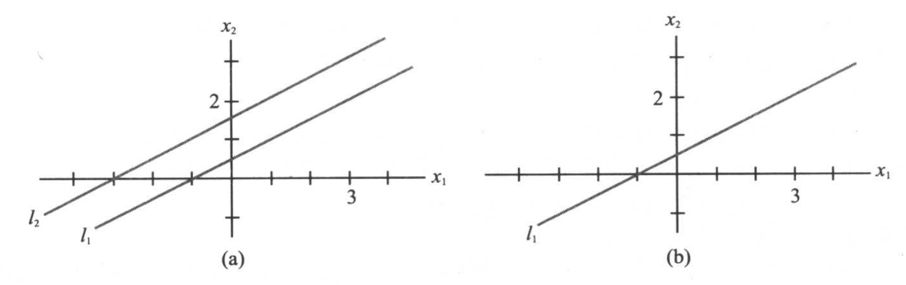


图 1-2 (a)无解；(b)无穷多解

图 1-1 和图 1-2 说明了线性方程组的下列一般事实，我们将在 1.2 节中加以验证.

一个线性方程组的解有下列三种情况：

1. 无解； 2. 有唯一解； 3. 有无穷多解.

如果一个线性方程组有唯一解或无穷多解，则称它是**相容**(consistent)的；如果无解，则称它是**不相容**(inconsistent)的.

1.1.1 矩阵记号

线性方程组的主要信息可以紧凑地记录在一个矩形阵列里，称之为**矩阵**(matrix). 已知方程组

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned} \tag{3}$$

其中每个变量的系数对齐成一列，则矩阵

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix}$$

称为方程组(3)的**系数矩阵**(coefficient matrix 或 matrix of coefficients)，矩阵

4

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix} \qquad (4)$$

称为方程组的**增广矩阵**(augmented matrix)(第2行中有一个0是因为第二个方程可以写成$0x_1+2x_2-8x_3=8$). 线性方程组的增广矩阵由系数矩阵加上方程组右端常数列组成.

矩阵的**维度**(size)是指矩阵所包含的行、列数. 增广矩阵(4)包含3行、4列，称为3×4(读作"3乘4")矩阵. 如果m和n是正整数，则$\boldsymbol{m\times n}$**矩阵**($m\times n$ matrix)是含有m行、n列的矩形数值阵列(行数通常写在前面). 在下面的例题中，矩阵记号将简化计算.

1.1.2　解线性方程组

本节与下一节介绍的是求解线性方程组的一套系统的算法. 算法的基本策略是*把方程组替换为另一个更容易求解的等价的(与原方程组有相同解集的)方程组.*

粗略地说，就是用方程组第一个方程中的x_1项消去其他方程中的x_1项，然后用第二个方程中的x_2项消去其他方程中的x_2项，依此类推，直到最后得到一个非常简单的等价方程组.

化简线性方程组的三种基本变换是：把一个方程替换成它本身与另一个方程的倍数之和；交换两个方程；将方程的每一项乘以一个非零常数. 在例题1之后，我们将解释为什么这三种变换不改变方程组的解集.

【例题1】　解线性方程组(3).

解：下面采用矩阵记号与非矩阵记号分别描述消去过程，并将结果并排呈现，以便对照：

$$\begin{aligned} x_1-2x_2+x_3&=0 \\ 2x_2-8x_3&=8 \\ -4x_1+5x_2+9x_3&=-9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

保留第一个方程中的x_1，并且消去其余方程中的x_1. 为做到这一点，把方程1的4倍加到方程3上. 适当练习以后，这类计算通常可以用心算完成：

$$\begin{array}{rr} 4\times[\text{方程}1]: & 4x_1-8x_2+4x_3=0 \\ +[\text{方程}3]: & -4x_1+5x_2+9x_3=-9 \\ \hline [\text{新方程}3]: & -3x_2+13x_3=-9 \end{array}$$

5

用这一计算结果代替原方程组的方程3：

$$\begin{aligned} x_1-2x_2+x_3&=0 \\ 2x_2-8x_3&=8 \\ -3x_2+13x_3&=-9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

现在，将方程2乘以1/2，使得x_2的系数变成1.（这个计算将简化下一步的计算.）

$$\begin{aligned} x_1-2x_2+x_3&=0 \\ x_2-4x_3&=4 \\ -3x_2+13x_3&=-9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

用方程 2 中的 x_2消去方程 3 中的 $-3x_2$. 心算过程为:

$$\begin{array}{rr} 3\times[\text{方程 }2]: & 3x_2-12x_3=12 \\ +[\text{方程 }3]: & -3x_2+13x_3=-9 \\ \hline [\text{新方程 }3]: & x_3=3 \end{array}$$

新的方程组具有三角形形式:[1]

$$\begin{aligned} x_1-2x_2+x_3&=0 \\ x_2-4x_3&=4 \\ x_3&=3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

我们希望最终能消去方程 1 中的 $-2x_2$项, 不过, 先用方程 3 中的 x_3消去方程 2 中的 $-4x_3$以及方程 1 中的 $+x_3$, 效率会更高. 这两个心算过程为:

$$\begin{array}{rr} 4\times[\text{方程 }3]: & 4x_3=12 \\ +[\text{方程 }2]: & x_2-4x_3=4 \\ \hline [\text{新方程 }2]: & x_2\qquad=16 \end{array} \qquad \begin{array}{rr} -1\times[\text{方程 }3]: & -x_3=-3 \\ +[\text{方程 }1]: & x_1-2x_2+x_3=0 \\ \hline [\text{新方程 }1]: & x_1-2x_2\qquad=-3 \end{array}$$

易将以上两步变换的结果综合如下:

$$\begin{aligned} x_1-2x_2\qquad&=-3 \\ x_2\qquad&=16 \\ x_3&=3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

我们已经把方程 3 中 x_3项所在列上方的各项消去了, 现在回到方程 2 的 x_2项, 用它消去上方中的 $-2x_2$. 由于之前已经对 x_3进行了处理, 所以现在的计算并不涉及 x_3. 把方程 2 的两倍加到方程 1 上, 得到如下方程组:

6

$$\begin{cases} x_1\qquad\qquad=29 \\ \quad x_2\qquad=16 \\ \qquad x_3=3 \end{cases} \qquad \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

主要的工作已经完成. 可以看出原方程组有唯一解(29,16,3)(见图 1-3). 不过, 由于整个过程中计算量相当大, 我们最好养成验证的习惯. 为了验证(29,16,3)是一个解, 把这些值代入原方程组的左端, 计算得:

$$\begin{aligned} (29)-2(16)+(3)&=29-32+3=0 \\ 2(16)-8(3)&=32-24=8 \\ -4(29)+5(16)+9(3)&=-116+80+27=-9 \end{aligned}$$

得到的结果和原方程组右端一致, 所以(29,16,3)是原方程组的一个解. ■

例题 1 呈现了线性方程组方程上的变换与增广矩阵行上的变换之间的对应关系. 前面列出的三种基本变换对应于增广矩阵的下列变换.

初等行变换

1. (替换)用某行与另一行的数量倍数之和替换该行.[2]

---

1. 下一节将用准确的概念代替"三角形"这一直观描述.
2. 行替换的一个常见解释是"把另一行的倍数加到某一行上".

2. (交换)交换两行.

3. (数乘)将某行全体元素都乘以某一非零常数.

行变换可以应用于任意矩阵，而不仅仅是线性方程组的增广矩阵. 两个矩阵如果可以通过一系列的初等行变换进行转化，我们称这两个矩阵**行等价**(row equivalent).

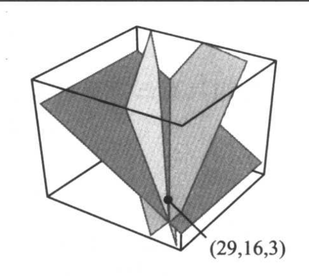


图1-3　原方程组中的每个方程都确定了3维空间中的一个平面. 点(29,16,3)同时落在这三个平面上

需要注意行变换是*可逆的*，这一点非常重要. 如果两行交换，那么再进行一次交换，它们可以回到原来的位置. 如果一行通过乘以非零常量 c 进行缩放，那么在得到的新行上乘以 $1/c$ 就得到原来的行. 最后，考虑两行的替换变换，比如说第1，2两行，假设第1行的 c 倍加到第2行上得到新的第2行. 这个变换的逆就是把第1行的 $-c$ 倍加到新的第2行上，其结果是原来的第2行. 见本节末尾习题29~32.

目前，我们只对线性方程组增广矩阵的行变换感兴趣. 假设一个方程组经行变换变为一个新的方程组，通过考虑各种形式的行变换，你能看到原方程组的任意解仍然是新方程组的解. 反过来，原方程组也可以由新方程组通过行变换得到，所以新方程组的任意解也是原方程组的解. 这就证明了以下事实：

7

如果两个线性方程组的增广矩阵是行等价的，那么这两个线性方程组有相同的解集.

尽管例题1很长，但稍加练习后你会发现计算起来其实很快. 本书正文及习题中的行变换通常都非常简单，因而你可以集中精力去理解概念及其思想. 尽管如此，你仍然必须学会如何正确地进行行变换，因为它在全书中都要用到.

本节余下的内容将要介绍，怎样在不完全求解线性方程组的情况下，利用行变换确定解集的大小.

1.1.3　存在性和唯一性问题

在1.2节中，我们将会看到为什么一个线性方程组或无解，或有唯一解，或有无穷解. 给定某个方程组，为了确定它属于哪种情形，我们问两个问题.

线性方程组的两个基本问题

1. 方程组是否相容，即是否*至少存在一个*解?
2. 如果解存在，解是否*仅有一个*，即解是否唯一?

这两个问题将以不同形式出现在整本书里. 在这一节和下一节中，我们将演示如何利用增广矩阵的行变换来回答这两个问题.

【例题2】　判断下列方程组是否相容：

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

解：这是例题1中的方程组. 假设我们已经进行了必要的行变换，得到三角形形式

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ x_3 &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

这时我们已经求出 x_3. 如果把 x_3 的值带入方程 2 中，我们可以计算出 x_2 的值，因此也可以从方程 1 中确定 x_1 的值. 所以解是存在的，方程组相容(事实上，因为 x_3 的取值唯一，所以 x_2 由方程 2 唯一确定，x_1 也由方程 1 唯一确定. 所以解是唯一的). ■

8

【例题 3】 判断下列方程组是否相容：

$$\begin{aligned} x_2 - 4x_3 &= 8 \\ 2x_1 - 3x_2 + 2x_3 &= 1 \\ 5x_1 - 8x_2 + 7x_3 &= 1 \end{aligned} \tag{5}$$

解： 增广矩阵为

$$\begin{bmatrix} 0 & 1 & -4 & 8 \\ 2 & -3 & 2 & 1 \\ 5 & -8 & 7 & 1 \end{bmatrix}$$

为了使方程 1 含有 x_1，将第 1，2 两行互换：

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 5 & -8 & 7 & 1 \end{bmatrix}$$

为了消去方程 3 中 $5x_1$ 这一项，把第 1 行乘以 $-5/2$ 加到第 3 行上：

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & -1/2 & 2 & -3/2 \end{bmatrix} \tag{6}$$

接下来，用方程 2 中的 x_2 消去方程 3 中的 $-(1/2)x_2$ 项. 把第 2 行乘以 1/2 加到第 3 行上：

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \end{bmatrix} \tag{7}$$

增广矩阵现在变成了三角形形式，为正确解释其含义，将它写为方程的形式：

$$\begin{aligned} 2x_1 - 3x_2 + 2x_3 &= 1 \\ x_2 - 4x_3 &= 8 \\ 0 &= 5/2 \end{aligned} \tag{8}$$

方程 $0 = 5/2$ 是 $0x_1 + 0x_2 + 0x_3 = 5/2$ 的简写形式. 这个三角形形式的方程组内明显存在矛盾. 由于 $0 = 5/2$ 恒假，所以不存在 x_1，x_2 和 x_3 的取值能满足方程组(8). 又因为方程组(8)和(5)有相同的解集，所以原方程组不相容(即无解)(见图 1-4). ■

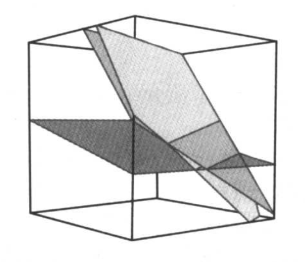


图 1-4 三个平面没有公共点，因此方程组不相容

9

特别注意(7)中的增广矩阵，其末行在三角形形式的不相容方程组中十分典型.

注记

在实际问题中，通常利用计算机来求解线性方程组. 对于系数方阵，计算机程序使用的几乎都是本节及1.2节所介绍的消元法，只是为提高精度稍作修改.

大多数商业和工业上的线性代数问题都是由采用浮点算术的计算机程序解决的. 在浮点算术中，数值采用十进制表示：$\pm 0.d_1\cdots d_p \times 10^r$，其中 r 为整数，p 是小数点右边的位数，通常在8到16之间. 对这类数值进行算术运算时，为了使其结果保留指定的存储位数，需要进行舍入(或截断)，因此结果往往不精确. 还有一种情形也会引起“舍入误差”，比如将1/3输入计算机，这是因为我们只能用某个有限长的浮点数近似地表示它对应的小数. 庆幸的是，浮点算术中的误差很少引发大的问题. 对于这一类需要注意的问题，本书的注记将会适时予以提醒.

基础练习

在完成本书的习题之前，应该先尝试基础练习. 基础练习的答案在每节习题后给出.

1. 口述为了求解方程组，应该执行的下一个基本行变换：[(a)可能有多个答案.]

a. $$\begin{aligned} x_1 + 4x_2 - 2x_3 + 8x_4 &= 12 \\ x_2 - 7x_3 + 2x_4 &= -4 \\ 5x_3 - x_4 &= 7 \\ x_3 + 3x_4 &= -5 \end{aligned}$$

b. $$\begin{aligned} x_1 - 3x_2 + 5x_3 - 2x_4 &= 0 \\ x_2 + 8x_3 \qquad &= -4 \\ 2x_3 \qquad &= 3 \\ x_4 &= 1 \end{aligned}$$

2. 一个线性方程组的增广矩阵经过初等行变换变成以下形式. 判断该方程组是否相容.

$$\begin{bmatrix} 1 & 5 & 2 & -6 \\ 0 & 4 & -7 & 2 \\ 0 & 0 & 5 & 0 \end{bmatrix}$$

3. (3,4,-2)是否是下列方程组的一个解？

$$\begin{aligned} 5x_1 - x_2 + 2x_3 &= 7 \\ -2x_1 + 6x_2 + 9x_3 &= 0 \\ -7x_1 + 5x_2 - 3x_3 &= -7 \end{aligned}$$

4. h，k 取什么值的时候，下列方程组相容？

$$\begin{aligned} 2x_1 - x_2 &= h \\ -6x_1 + 3x_2 &= k \end{aligned}$$

10

习题 1.1

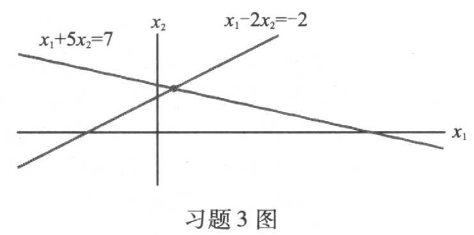
依照本节所介绍的系统消去法，对方程或增广矩阵运用初等行变换，求解习题1~4中的方程组.

1. $$\begin{aligned} x_1 + 5x_2 &= 7 \\ -2x_1 - 7x_2 &= -5 \end{aligned}$$

2. $$\begin{aligned} 2x_1 + 4x_2 &= -4 \\ 5x_1 + 7x_2 &= 11 \end{aligned}$$

3. 找出既在直线 $x_1 + 5x_2 = 7$ 上也在直线 $x_1 - 2x_2 = -2$ 上的点 (x_1, x_2). 如图所示.

4. 找出直线 $x_1 - 5x_2 = 1$ 和 $3x_1 - 7x_2 = 5$ 的交点.



习题3图

把习题5和6中的每个矩阵看作是方程组的增广矩阵. 口述：求解方程组的过程中，

接下来要执行的两步初等行变换.

5. $\begin{bmatrix} 1 & -4 & 5 & 0 & 7 \\ 0 & 1 & -3 & 0 & 6 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & -5 \end{bmatrix}$

6. $\begin{bmatrix} 1 & -6 & 4 & 0 & -1 \\ 0 & 2 & -7 & 0 & 4 \\ 0 & 0 & 1 & 2 & -3 \\ 0 & 0 & 3 & 1 & 6 \end{bmatrix}$

在习题 7 ~ 10 中，线性方程组的增广矩阵已经化简成题中所示的形式. 对每道题目，继续进行适当的行变换，并且给出原方程组的解集.

7. $\begin{bmatrix} 1 & 7 & 3 & -4 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix}$

8. $\begin{bmatrix} 1 & -4 & 9 & 0 \\ 0 & 1 & 7 & 0 \\ 0 & 0 & 2 & 0 \end{bmatrix}$

9. $\begin{bmatrix} 1 & -1 & 0 & 0 & -4 \\ 0 & 1 & -3 & 0 & -7 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 0 & 0 & 2 & 4 \end{bmatrix}$

10. $\begin{bmatrix} 1 & -2 & 0 & 3 & -2 \\ 0 & 1 & 0 & -4 & 7 \\ 0 & 0 & 1 & 0 & 6 \\ 0 & 0 & 0 & 1 & -3 \end{bmatrix}$

求解习题 11 ~ 14 中的方程组.

11. $\begin{aligned} x_2 + 4x_3 &= -5 \\ x_1 + 3x_2 + 5x_3 &= -2 \\ 3x_1 + 7x_2 + 7x_3 &= 6 \end{aligned}$

12. $\begin{aligned} x_1 - 3x_2 + 4x_3 &= -4 \\ 3x_1 - 7x_2 + 7x_3 &= -8 \\ -4x_1 + 6x_2 - x_3 &= 7 \end{aligned}$

13. $\begin{aligned} x_1 - 3x_3 &= 8 \\ 2x_1 + 2x_2 + 9x_3 &= 7 \\ x_2 + 5x_3 &= -2 \end{aligned}$

14. $\begin{aligned} x_1 - 3x_2 &= 5 \\ -x_1 + x_2 + 5x_3 &= 2 \\ x_2 + x_3 &= 0 \end{aligned}$

判断习题 15 和 16 中的方程组是否相容.

15. $\begin{aligned} x_1 + 3x_3 &= 2 \\ x_2 - 3x_4 &= 3 \\ -2x_2 + 3x_3 + 2x_4 &= 1 \\ 3x_1 + 7x_4 &= -5 \end{aligned}$

16. $\begin{aligned} x_1 - 2x_4 &= -3 \\ 2x_2 + 2x_3 &= 0 \\ x_3 + 3x_4 &= 1 \\ -2x_1 + 3x_2 + 2x_3 + x_4 &= 5 \end{aligned}$

17. $x_1 - 4x_2 = 1$，$2x_1 - x_2 = -3$ 和 $-x_1 - 3x_2 = 4$ 这三条直线是否相交于一点？请给出解释.

18. $x_1 + 2x_2 + x_3 = 4$，$x_2 - x_3 = 1$ 和 $x_1 + 3x_2 = 0$ 这三个平面是否至少相交于一点？请给出解释.

在习题 19 ~ 22 中，确定 h 的值，使题中所示矩阵为相容线性方程组的增广矩阵.

11

19. $\begin{bmatrix} 1 & h & 4 \\ 3 & 6 & 8 \end{bmatrix}$　20. $\begin{bmatrix} 1 & h & -3 \\ -2 & 4 & 6 \end{bmatrix}$　21. $\begin{bmatrix} 1 & 3 & -2 \\ -4 & h & 8 \end{bmatrix}$　22. $\begin{bmatrix} 2 & -3 & h \\ -6 & 9 & 5 \end{bmatrix}$

在习题 23 和 24 中，本节的一些关键命题或被直接引用，或被重新陈述(仍然正确)，或被改动，使其在一些情况下错误. 判断各命题的真假，并说明理由(如果命题为真，指出类似的陈述大致出现在书中的哪个位置，或者参考的是哪个定义或定理. 如果命题为假，指出被错误引用或使用的命题所在的位置，或者举例说明这个陈述并不是在所有情况下成立). 类似的真假判断问题还会在本书的其他章节中出现.

23. a. 所有的初等行变换都是可逆的.

b. 一个 5×6 矩阵有 6 行.

c. 一个含有变量 x_1，x_2，…，x_n的线性方程组的解集是一列数$(s_1, s_2, \cdots, s_n)$，使得用 s_1，$s_2, \cdots, s_n$分别替换方程中的 $x_1, x_2, \cdots, x_n$后，方程组中的每个方程都成立.

d. 有关线性方程组的两个基本问题是存在性和唯一性.

24. a. 增广矩阵的初等行变换不会改变相对应线性方程组的解集.
 b. 两个矩阵如果有相同的行数，则行等价.
 c. 一个相容方程组不止有一个解.
 d. 如果两个线性方程组有相同的解集，则等价.
25. 找出含有 g，h，k 的方程，使得下列增广矩阵对应于一个相容的方程组.

$$\begin{bmatrix} 1 & -4 & 7 & g \\ 0 & 3 & -5 & h \\ -2 & 5 & -9 & k \end{bmatrix}$$

26. 构造三个不同的增广矩阵，使其对应于解集为 $x_1=-2$，$x_2=1$，$x_3=0$ 的线性方程组.
27. 假设以下方程组对于 f，g 的任意取值都相容，那么关于系数 c 和 d，你能得出什么结论？说明理由.

$$\begin{aligned} x_1 + 3x_2 &= f \\ cx_1 + dx_2 &= g \end{aligned}$$

28. 假设 a，b，c，d 为常数，a 不为零，且下面方程组对于 f，g 的任意取值都相容. 那么关于系数 a，b，c，d，你能得出什么结论？说明理由.

$$\begin{aligned} ax_1 + bx_2 &= f \\ cx_1 + dx_2 &= g \end{aligned}$$

在习题 29 ~ 32 中，求将第一个矩阵变换成第二个矩阵的初等行变换，以及将第二个矩阵变换成第一个矩阵的逆初等行变换.

29. $\begin{bmatrix} 0 & -2 & 5 \\ 1 & 4 & -7 \\ 3 & -1 & 6 \end{bmatrix}$, $\begin{bmatrix} 1 & 4 & -7 \\ 0 & -2 & 5 \\ 3 & -1 & 6 \end{bmatrix}$

30. $\begin{bmatrix} 1 & 3 & -4 \\ 0 & -2 & 6 \\ 0 & -5 & 9 \end{bmatrix}$, $\begin{bmatrix} 1 & 3 & -4 \\ 0 & 1 & -3 \\ 0 & -5 & 9 \end{bmatrix}$

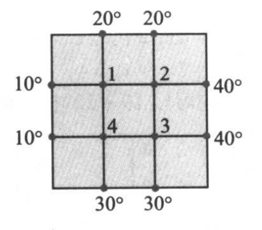
31. $\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 5 & -2 & 8 \\ 4 & -1 & 3 & -6 \end{bmatrix}$, $\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 5 & -2 & 8 \\ 0 & 7 & -1 & -6 \end{bmatrix}$

32. $\begin{bmatrix} 1 & 2 & -5 & 0 \\ 0 & 1 & -3 & -2 \\ 0 & -3 & 9 & 5 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 & -5 & 0 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 0 & -1 \end{bmatrix}$

热传导研究中的一个重要问题是，已知金属薄片边界附近的温度，确定其上稳态温度的分布. 假设下图所示的金属片表示一根金属柱的横截面，并且忽略与盘片垂直方向上的热量传递. 设 $T_1,\cdots,T_4$ 表示图中 4 个内部网格节点的温度. 一个节点的温度约等于四个相邻节点(上、下、左、右)的平均温度.[1] 例如，

12

$$T_1=(10+20+T_2+T_4)/4, \quad 或 \quad 4T_1-T_2-T_4=30$$



习题 33、34 图

33. 写出一个含有 4 个方程的方程组，该方程组的解给出对温度 $T_1,\cdots,T_4$ 的估计.
34. 求解习题 33 中的方程组. [提示：为加快计算速度，可以在执行替换变换前先交换第 1 行和第 4 行.]

基础练习答案

1. a. 对于“手算”来说，最佳的选择是交换方程 3 和 4. 另一个可能是将方程 3 乘以 1/5. 或

1. 参考 Frank M. White，*Heat and Mass Transfer*（Reading，MA：Addison-Wesley Publishing. 1991）pp. 145 ~ 149.

者把方程 4 替换成它自身加上方程 3 乘以 $-1/5$ 的积. 在任何情况下，都不要用方程 2 中的 x_2 消去方程 1 中的 $4x_2$. 直到我们得到三角形形式的方程组，并且 x_3 和 x_4 从前两个方程中被消去.

b. 该方程组为三角形形式. 进一步的化简开始于方程 4 中的 x_4. 用这个 x_4 消去前面所有方程中的 x_4. 步骤就是把方程 4 的 2 倍加到方程 1 上. 然后处理方程 3，将它乘以 1/2，用得到的新方程消去前面方程中的 x_3.

2. 对应于增广矩阵的方程组是

$$\begin{aligned} x_1 + 5x_2 + 2x_3 &= -6 \\ 4x_2 - 7x_3 &= 2 \\ 5x_3 &= 0 \end{aligned}$$

由第三个方程得到 $x_3=0$，这显然是 x_3 的一个可行解. 在消去方程 1 和方程 2 中的 x_3 后，可以继续求出 x_1，x_2 的唯一值. 因此解是存在的，并且唯一. 可以将本题与例题 3 做对比.

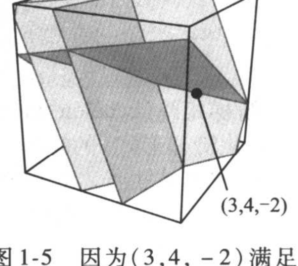


图 1-5 因为(3,4，−2)满足前两个方程，所以它在前两个平面的交线上. 但(3,4，−2)不同时满足三个方程，因此它并不同在三个平面上

3. 很容易检验某列指定的数值是否为解（见图 1-5）. 令 $x_1=3$，$x_2=4$，$x_3=-2$，得到

$$\begin{aligned} 5(3)-(4)+2(-2) &= 15-4-4=7 \\ -2(3)+6(4)+9(-2) &= -6+24-18=0 \\ -7(3)+5(4)-3(-2) &= -21+20+6=5 \end{aligned}$$

尽管满足前两个方程，但第三个方程不成立，所以(3,4，−2)不是方程组的解. 注意，在进行变量代换时使用了圆括号，为避免计算错误，我们强烈推荐这样做.

4. 当第二个方程被替换为它自身和第一个方程的 3 倍的和，方程组变为

$$\begin{aligned} 2x_1 - x_2 &= h \\ 0 &= k+3h \end{aligned}$$

如果 $k+3h$ 不为零，则方程组无解. 对于使 $k+3h=0$ 成立的 h 和 k 的任意取值，方程组都相容.

13

1.2 行化简和阶梯形式

本节中，我们将 1.1 节中的方法提炼成能够分析任意线性方程组的行化简算法.[1]仅仅利用这个算法的第一部分，我们就可以回答 1.1 节中提出的存在性与唯一性这两个基本问题.

这个算法适用于任意矩阵，不论这个矩阵是否被看作是线性方程组的增广矩阵. 因此本节的第一部分考虑任意矩阵. 我们首先介绍包括 1.1 节“三角形”矩阵在内的两类重要矩阵. 在下面的定义中，矩阵的非零行或非零列是指至少包含一个非零元素的行或列；行的**首项元素**(leading entry)是指(非零行中)最左边的非零元素.

---

1. 我们所用的算法是通常被称为高斯消去法的一个变形. 公元前 250 年，中国数学家就已经使用了针对线性方程组的一种类似的消去法. 但这一算法，直到 19 世纪被著名的德国数学家卡尔·弗里德里希·高斯发现，才为西方世界所知. 1888 年，德国工程师 Wilhelm Jordan 将其收录于一本测地学著作，这一算法因而得以推广.

【定义】 如果一个矩阵满足以下三个性质，则称该矩阵具有**阶梯形式**(echelon form)或**行阶梯形式**(row echelon form)：

1. 所有的非零行均在任一全为零的行之上.
2. 每一行首项元素所在的列，都在上一行首项元素所在的列的右边.
3. 同一列中位于首项元素下方的所有元素均为零.

如果一个具有阶梯形式的矩阵还满足以下条件，则称该矩阵具有**简化阶梯形式**(reduced echelon form)或**行简化阶梯形式**(reduced row echelon form)：

4. 非零行中的首项元素均为 1.
5. 每个首项元素 1 在其所处的列中是唯一的非零元素.

阶梯矩阵(echelon matrix)[**简化阶梯矩阵**(reduced echelon matrix)]是具有阶梯形式(简化阶梯形式)的矩阵. 性质 2 是指全体首项元素排成了指向矩阵右下角的阶梯状. 性质 3 是性质 2 的一个简单推论，但为了强调我们仍然保留.

1.1 节中的“三角形”矩阵，比如

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \end{bmatrix} \text{ 和 } \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

14

都是具有阶梯形式的矩阵. 事实上，第二个矩阵具有行简化阶梯形式. 下面还有一些例子.

【例题 1】 下面的矩阵都具有阶梯形式. 首项元素 [■] 可以取任意非零值；星号 [*]标记的元素可以取任意值(包括0).

$$\begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * \end{bmatrix}$$

下面的矩阵具有行简化阶梯形式，因为其首项元素均为 1，并且首项元素 1 的上、下方元素均为 0.

$$\begin{bmatrix} 1 & 0 & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & * & 0 & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 1 & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \end{bmatrix}$$

■

使用不同的行化简序列，任意非零矩阵都可以**行化简**(row reduced，即用初等行变换进行变换)成多个阶梯矩阵. 然而，从一个矩阵出发得到的简化阶梯形式矩阵是唯一的. 下面的定理将在本书末尾附录 A 中予以证明.

【定理 1】 **简化阶梯形式的唯一性**

每个矩阵都与唯一一个简化阶梯矩阵行等价.

如果矩阵 $\boldsymbol{A}$ 与一个阶梯矩阵 $\boldsymbol{U}$ 行等价，我们称 $\boldsymbol{U}$ 为 **$\boldsymbol{A}$ 的阶梯形式**(echelon form of $\boldsymbol{A}$)或**行阶梯形式**(row echelon form)；如果 $\boldsymbol{U}$ 具有简化阶梯形式，则称其为 **$\boldsymbol{A}$ 的简化阶梯形式**(reduced echelon form of $\boldsymbol{A}$). [大多数能够处理矩阵的矩阵软件和计算器将(行)简化阶梯形式缩写为 RREF. 有的将(行)阶梯形式缩写为 REF.]

1.2.1 主元位置

当矩阵进行行变换得到阶梯形式后，再进一步行变换得到简化阶梯形式，这个过程并没有改变首项元素的位置. 由简化阶梯形式的唯一性可知，在由给定矩阵得到的任意阶梯矩阵中，首项元素的位置相同. 这些首项元素对应于简化阶梯矩阵中的首项元素 1.

15

【定义】 **矩阵 $\boldsymbol{A}$ 中的主元位置(pivot position)是 $\boldsymbol{A}$ 的简化阶梯形式中首项元素 1 所对应的位置. 主元列(pivot column)是 $\boldsymbol{A}$ 中包含主元位置的列.**

在例题 1 中，正方形 [■] 标识了主元位置. 前四章的许多基本概念都与矩阵的主元位置有联系.

【例题 2】 将下列矩阵 $\boldsymbol{A}$ 化简为阶梯形式，并且求 $\boldsymbol{A}$ 的主元列.

$$A=\begin{bmatrix}0&-3&-6&4&9\\-1&-2&-1&3&1\\-2&-3&0&3&-1\\1&4&5&-9&-7\end{bmatrix}$$

解：使用1.1 节中的基本策略. 最左边非零列的第一个位置即是第一个主元位置，这里必须放上一个非零元或主元. 一个好的办法是互换第 1，4 两行(这样下面的心算就不涉及分数运算).

$$\begin{bmatrix}\overset{\text{主元}}{1}&4&5&-9&-7\\-1&-2&-1&3&1\\-2&-3&0&3&-1\\\underset{\uparrow\,\text{主元列}}{0}&-3&-6&4&9\end{bmatrix}$$

将第 1 行的倍数加到下面的行上，使主元 1 下方元素变成 0，这就得到下面的矩阵(1)第 2 行的主元位置必须尽可能地靠左，即位于第 2 列. 我们选择在这个位置上的 2 作为下一个主元.

$$\begin{bmatrix}1&4&5&-9&-7\\0&\overset{\text{主元}}{2}&4&-6&-6\\0&5&10&-15&-15\\0&\underset{\uparrow\,\text{下一个主元列}}{-3}&-6&4&9\end{bmatrix} \tag{1}$$

把第 2 行乘以 $-5/2$ 加到第 3 行上，再把第 2 行乘以 $3/2$ 加到第 4 行上. 可得，

$$\begin{bmatrix}1&4&5&-9&-7\\0&2&4&-6&-6\\0&0&0&0&0\\0&0&0&-5&0\end{bmatrix} \tag{2}$$

16

(2)中的矩阵与1.1节中遇到的矩阵有所不同. 在这里无法生成第3列的主元(我们不能使用第1行或者第2行，因为这样做会破坏已经产生的首项元素的阶梯分布). 如果我们互换第3行和第4行，我们就能得到第4列的主元.

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{一般形式:} \quad \begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

主元　主元列

该矩阵具有阶梯形式，同时也表明矩阵 A 的主元列是第1，2，4列.

$$A = \begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix} \tag{3}$$

主元位置　主元列

■

如例题2所示，**主元**(pivot)是位于主元位置的非零数，必要时可以经由行变换用它们生成零. 例题2中的主元是1，2和-5. 注意，(3)中突出显示了 A 的主元位置，这些位置上的元素并不是1，2和-5. 事实上，不同的行变换序列可能生成不同的主元集合. 此外，如果进行行缩放把主元变成1，则主元在阶梯形式中将不出现(这将方便进行手算).

以例题2为指导，我们已经给出了将矩阵化为阶梯矩阵或简化阶梯矩阵的一种有效方法. 认真学习并掌握这种方法，将使你在后面的课程中受益良多.

1.2.2　行化简算法

下面的算法由四个步骤组成，最后得到阶梯矩阵. 第5个步骤将得到简化阶梯矩阵. 我们用例题3来演示这个算法.

【例题3】　应用初等行变换，先将下列矩阵转化为阶梯形式，然后再转化为简化阶梯形式：

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

17

解：

第1步
　从最左边的非零列开始. 这是一个主元列，主元位置在顶端.

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

主元列

> **第 2 步**
> 选择主元列中的一个非零元素为主元，如有必要，可通过行交换将这个元素移到主元位置上.

互换第 1 行和第 3 行(也可以互换第 1 行和第 2 行).

$$\begin{bmatrix} \overset{\text{主元}}{3} & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

> **第 3 步**
> 使用行替换变换来生成主元下方所有的零元.

作为一个准备步骤，我们可以把最上面的一行除以主元 3. 但是第 1 列中只有两个 3，容易看出只需将第 1 行乘以 −1 加到第 2 行上.

$$\begin{bmatrix} \overset{\text{主元}}{3} & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

> **第 4 步**
> 盖住（或忽略不管）含有主元位置的行和它上面所有的行. 对余下的子矩阵应用第 1 步到第 3 步. 重复这个过程，直到再也没有非零行需要修改.

盖住第 1 行，第 1 步显示列 2 是下一个主元列；对于第 2 步，我们将选择这一列顶端的元素为主元.

18

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & \overset{\text{主元}}{2} & -4 & 4 & 2 & -6 \\ 0 & \underset{\text{下一个主元列}}{3} & -6 & 6 & 4 & -5 \end{bmatrix}$$

对于第 3 步，我们可以选择先将子矩阵的最上面一行除以主元 2. 不过在这里，我们直接把最上面一行乘以 −3/2 加到下一行上. 这就得到：

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

现在应用第 4 步，当我们盖住包含第二个主元位置的行时，新的子矩阵仅含有一行：

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & \underset{\text{主元}}{1} & 4 \end{bmatrix}$$

不需要在子矩阵上进行第 1 步到第 3 步的操作，我们已经得到整个矩阵的阶梯形式. 如果要得到简化阶梯形式，则需多执行一步.

第5步

从最右边的主元开始，顺次向左，将每个主元上方的元素变成0. 如果主元不是1，则用数乘变换将它变成1.

最右边的主元在第3行. 把第3行的适当倍数加到第2行和第1行上，生成主元上方的0.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} \leftarrow \text{第1行} + (-6) \cdot \text{第3行} \\ \leftarrow \text{第2行} + (-2) \cdot \text{第3行} \\ \ \end{matrix}$$

下一个主元在第2行. 除以主元以缩放该行.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \quad \leftarrow \text{用} 1/2 \text{进行行缩放}$$

把第2行乘以9加到第1行上，生成第2列中的0.

$$\begin{bmatrix} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} \leftarrow \text{第1行} + (9) \cdot \text{第2行} \\ \ \\ \ \end{matrix}$$

19

最后，除以主元3以缩放第1行.

$$\begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \quad \leftarrow \text{用} 1/3 \text{进行行缩放}$$

这即是原矩阵的简化阶梯形式. ■

第1步到第4步合称行化简算法的**正向阶段**(forward phase)，生成唯一简化阶梯形式的第5步称为**反向阶段**(backward phase).

注记

在上面的第2步中，计算机程序通常选择一列中绝对值最大的元素为主元. 使用这个称作**部分主元法**(partial pivoting)的方法，可以减少计算过程中的舍入误差.

1.2.3 线性方程组的解

上述行化简算法应用到方程组的增广矩阵上时，将直接得出线性方程组解集的显式表示.

例如，假设线性方程组的增广矩阵已经变换成等价的简化阶梯形式

$$\begin{bmatrix} 1 & 0 & -5 & 1 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

增广矩阵有4列，因此有三个变量. 相应的方程组为

$$\begin{aligned} x_1 \quad -5x_3 &= 1 \\ x_2 + x_3 &= 4 \\ 0 &= 0 \end{aligned} \tag{4}$$

变量 x_1，x_2对应于矩阵中的主元列，称为**基本变量**(basic variable).[1]另一变量 x_3称为**自由变量**(free variable).

如果方程组是相容的，比如(4)，则只要解出简化方程组，将基本变量用自由变量表示出来，就得到方程组解集的显式表示. 因为基本变量在简化阶梯形式中仅出现在一个方程里，所以这个方法是可行的. 在(4)中，我们可以解第一个方程得到 x_1，解第二个方程得到 x_2. 忽略第三个方程，因为它对变量没有任何约束.

20

$$\begin{cases} x_1 = 1 + 5x_3 \\ x_2 = 4 - x_3 \\ x_3 \text{ 是自由变量} \end{cases} \tag{5}$$

称 x_3是"自由"的，意思是我们可以自由地选择 x_3的值. 一旦选定，(5)中的公式就可以确定 x_1，x_2的值. 例如，$x_3=0$时，解为(1,4,0)；$x_3=1$时，解为(6,3,1). x_3的每个不同的取值都确定方程组的一个(不同)解，方程组的每个解由一个 x_3的取值确定.

(5)中的解称为方程组的**通解**(general solution)，因为它给出了所有解的显式表示.

【例题4】 求出线性方程组的通解，其增广矩阵已经化简为：

$$\begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

解：矩阵为阶梯形式，但在求出基本变量之前，我们需要得到简化阶梯形式. 下面的步骤将完成行化简过程，矩阵前面的符号 ~ 表示该矩阵与前一个矩阵行等价.

$$\begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 2 & -5 & 0 & 10 \\ 0 & 0 & 2 & -8 & 0 & 10 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 6 & 2 & -5 & 0 & 10 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

增广矩阵有6列，因此有5个变量. 对应的方程组为

$$\begin{aligned} x_1 + 6x_2 \quad + 3x_4 \quad &= 0 \\ x_3 - 4x_4 \quad &= 5 \\ x_5 &= 7 \end{aligned} \tag{6}$$

矩阵的主元列是1，3，5列，所以基本变量是 x_1，x_3和 x_5. 余下的变量 x_2，x_4必定是自由变量. 求出基本变量，我们就得到通解

21

$$\begin{cases} x_1 = -6x_2 - 3x_4 \\ x_2 \text{ 是自由变量} \\ x_3 = 5 + 4x_4 \\ x_4 \text{ 是自由变量} \\ x_5 = 7 \end{cases} \tag{7}$$

1. 有的教材中称之为首项变量(leading variable)，因为它们对应的是含有首项元素的列.

注意，x_5的值已经被方程组(6)的第三个方程确定了. ■

1.2.4　解集的参数表示

通解(5)和(7)中的表示是解集的参数表示，自由变量作为参数出现在其中. 求解一个方程组就相当于找到解集的参数表示或者确定解集为空.

只要一个方程组是相容的，且有自由变量，则解集就可以有多种参数表示. 例如，在方程组(4)中，我们可以把方程2的5倍加到方程1上，得到等价方程组

$$\begin{aligned} x_1 + 5x_2 \quad &= 21 \\ x_2 + x_3 &= 4 \end{aligned}$$

我们可以把x_2看作参量，求出用x_2表示的x_1和x_3，我们也可以得到解集的精确表示. 然而，为了保持一致，我们约定总是把自由变量当作参数来表示一个解集. 本书最后部分的答案也遵循这个约定.

只要一个方程组是不相容的，就算方程组有自由变量，解集也为空集. 在这种情况下，解集没有参数表示.

1.2.5　矩阵的回代

考虑下列方程组，其增广矩阵是阶梯形式，但不是简化阶梯形式：

$$\begin{aligned} x_1 - 7x_2 + 2x_3 - 5x_4 + 8x_5 &= 10 \\ x_2 - 3x_3 + 3x_4 + x_5 &= -5 \\ x_4 - x_5 &= 4 \end{aligned}$$

计算机程序将使用矩阵回代方法求解该方程组，这个方法优于计算简化阶梯形式. 就是说，这个程序通过解方程3得出用x_5表示的x_4，并用这个表示替换方程2中的x_4，解方程2得到x_2，然后把x_2和x_4的表示替换到方程1中，求出x_1.

在行化简的反向阶段，我们用于生成简化阶梯形矩阵的矩阵形式化算法，其运算量与矩阵回代相同. 但矩阵形式化算法的规则很大程度上降低了手算过程中出错的可能性. 因此强烈建议大家只用简化阶梯形式求解方程组！本书附带的学习指南中对如何准确、快速地施行行变换提供了一些有用的建议.

22

注记

一般地，行化简的正向阶段比反向阶段长很多. 求解方程组的算法运算量通常用浮点运算来衡量. 一次**浮点运算**(flop)是指两个浮点实数上的一次算术运算(+，-，*，/).[1]对于一个$n\times(n+1)$矩阵，简化为阶梯矩阵需要进行$2n^3/3+n^2/2-7n/6$次浮点运算(当n取适当大的值，比如$n\geqslant 30$时，约为$2n^3$次浮点运算). 而进一步简化成简化阶梯形式需要至少n^2次浮点运算.

1.2.6　存在性和唯一性问题

尽管非简化阶梯形式并不是求解方程组的一个好工具，但它恰好能回答1.1节中提出的两个基本问题.

---

1. 浮点运算传统上仅仅指乘法或除法运算，这是因为加法和减法的执行时间非常短，可以忽略. 随着计算机体系结构的不断改进，现在，这里给出的浮点运算定义更可取. 见Golub and Van Loan, *Matrix Computations*, 2nd ed. (Baltimore：The Johns Hopkins Press, 1989), pp. 19～20.

【例题 5】 判断下列方程组解的存在性和唯一性：

$$\begin{aligned} 3x_2 - 6x_3 + 6x_4 + 4x_5 &= -5 \\ 3x_1 - 7x_2 + 8x_3 - 5x_4 + 8x_5 &= 9 \\ 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 &= 15 \end{aligned}$$

解： 例题 3 中将该方程组的增广矩阵行化简得到

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \tag{8}$$

基本变量是 x_1，x_2和 x_5；自由变量为 x_3和 x_4. 并且没有像 $0 = 1$ 这样导致方程组不相容的方程. 所以我们可以用矩阵回代方法来求解. 解的存在性在(8)中已经很清楚了. 此外，因为有三个自由变量，解并不唯一. x_3和 x_4的每种不同取值都确定了不同的解. 因此，该方程组有无穷多解. ■ 23

当一个方程组呈阶梯形式，且不含有形如 $0 = b$(b 不为零)的方程时，每个非零方程都包含一个系数不为零的基本变量. 此时，或者基本变量完全被确定(没有自由变量)，或者至少有一个基本变量可以用一个或多个自由变量表示. 前一种情况，解是唯一的；后一种情况，则有无穷多解(自由变量的每个取值对应一个解).

上述讨论证明了下面的定理.

【定理 2】 存在性和唯一性定理

一个线性方程组是相容的当且仅当增广矩阵的最右边一列不为主元列，即当且仅当增广矩阵的阶梯形式没有如下形式的行

$$[0 \cdots 0 \ b] \quad \text{其中 } b \text{ 不为零}$$

如果线性方程组相容，则(i)没有自由变量时，有唯一解，或(ii)至少有一个自由变量时，有无穷多解.

下面的步骤概括了如何求出并表示线性方程组的所有解.

使用行变换求解线性方程组

1. 写出方程组的增广矩阵.
2. 使用行变换法则得到等价的阶梯形式矩阵. 判定方程组是否相容. 如果无解，则停止；否则，进行下一步.
3. 继续进行行变换得到简化阶梯形式.
4. 写出第 3 步所得矩阵对应的方程组.
5. 改写第 4 步得到的每个非零方程，将其中的基本变量用自由变量来表示.

基础练习

1. 求线性方程组的通解，其增广矩阵为：

$$\begin{bmatrix} 1 & -3 & -5 & 0 \\ 0 & 1 & 1 & 3 \end{bmatrix}$$

2. 求下列线性方程组的通解： 24

$$\begin{aligned} x_1 - 2x_2 - x_3 + 3x_4 &= 0 \\ -2x_1 + 4x_2 + 5x_3 - 5x_4 &= 3 \\ 3x_1 - 6x_2 - 6x_3 + 8x_4 &= 2 \end{aligned}$$

习题　1.2

在习题1和2中，判断哪些矩阵是简化阶梯形式，哪些仅是阶梯形式.

1. a. $\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&1\end{bmatrix}$　b. $\begin{bmatrix}1&0&1&0\\0&1&1&0\\0&0&0&1\end{bmatrix}$　c. $\begin{bmatrix}1&0&0&0\\0&1&1&0\\0&0&0&0\\0&0&0&1\end{bmatrix}$　d. $\begin{bmatrix}1&1&0&1&1\\0&2&0&2&2\\0&0&0&3&3\\0&0&0&0&4\end{bmatrix}$

2. a. $\begin{bmatrix}1&1&0&1\\0&0&1&1\\0&0&0&0\end{bmatrix}$　b. $\begin{bmatrix}1&1&0&0\\0&1&1&0\\0&0&1&1\end{bmatrix}$　c. $\begin{bmatrix}1&0&0&0\\1&1&0&0\\0&1&1&0\\0&0&1&1\end{bmatrix}$　d. $\begin{bmatrix}0&1&1&1&1\\0&0&2&2&2\\0&0&0&0&3\\0&0&0&0&0\end{bmatrix}$

对习题3和4中的矩阵进行行化简，使之具有简化阶梯形式. 在原始矩阵和最后得到的矩阵中圈出主元位置，列出主元列.

3. $\begin{bmatrix}1&2&3&4\\4&5&6&7\\6&7&8&9\end{bmatrix}$　　4. $\begin{bmatrix}1&3&5&7\\3&5&7&9\\5&7&9&1\end{bmatrix}$

5. 描述 2×2 矩阵所有可能的阶梯形式. 使用例题1的第一部分那样的符号：■, $*$ 和0.
6. 对非零 3×2 矩阵重做习题5.

求习题7～14中增广矩阵所对应的方程组的通解.

7. $\begin{bmatrix}1&3&4&7\\3&9&7&6\end{bmatrix}$　8. $\begin{bmatrix}1&4&0&7\\2&7&0&10\end{bmatrix}$　9. $\begin{bmatrix}0&1&-6&5\\1&-2&7&-6\end{bmatrix}$　10. $\begin{bmatrix}1&-2&-1&3\\3&-6&-2&2\end{bmatrix}$

11. $\begin{bmatrix}3&-4&2&0\\-9&12&-6&0\\-6&8&-4&0\end{bmatrix}$　12. $\begin{bmatrix}1&-7&0&6&5\\0&0&1&-2&-3\\-1&7&-4&2&7\end{bmatrix}$

13. $\begin{bmatrix}1&-3&0&-1&0&-2\\0&1&0&0&-4&1\\0&0&0&1&9&4\\0&0&0&0&0&0\end{bmatrix}$　14. $\begin{bmatrix}1&2&-5&-6&0&-5\\0&1&-6&-3&0&2\\0&0&0&0&1&0\\0&0&0&0&0&0\end{bmatrix}$

习题15和16使用了例题1中阶梯矩阵的记号. 假设每个矩阵都代表了一个线性方程组的增广矩阵. 判断每个方程组是否相容. 如果方程组相容，判断它的解是否唯一.

25

15. a. $\begin{bmatrix}\blacksquare&*&*&*\\0&\blacksquare&*&*\\0&0&\blacksquare&0\end{bmatrix}$ b. $\begin{bmatrix}0&\blacksquare&*&*&*\\0&0&\blacksquare&*&*\\0&0&0&0&\blacksquare\end{bmatrix}$　16. a. $\begin{bmatrix}\blacksquare&*&*\\0&\blacksquare&*\\0&0&0\end{bmatrix}$ b. $\begin{bmatrix}\blacksquare&*&*&*&*\\0&0&\blacksquare&*&*\\0&0&0&\blacksquare&*\end{bmatrix}$

在习题17和18中，确定 h 的值，使得题中矩阵对应的线性方程组相容.

17. $\begin{bmatrix}2&3&h\\4&6&7\end{bmatrix}$　　18. $\begin{bmatrix}1&-3&-2\\5&h&-7\end{bmatrix}$

在习题19和20中，选择 h 和 k 的值，使得方程组(a)无解，(b)有唯一解，(c)有很多解. 分别给出每一部分的答案.

19. $x_1+hx_2=2$
　　$4x_1+8x_2=k$

20. $x_1+3x_2=2$
　　$3x_1+hx_2=k$

在习题21和22中，判断每个命题的真假，并说明理由.[1]

21. a. 在某些情况下，一个矩阵可以通过不同的行变换序列变成多个不同的简化阶梯矩阵.
    b. 行化简算法仅适用于线性方程组的增广矩阵.
    c. 线性方程组中的一个基本变量对应于系数矩阵中的一个主元列.
    d. 求一个线性方程组解集的参数表示等同于求解方程组.
    e. 如果增广矩阵中的一行为[0 0 0 5 0]，则相应的线性方程组不相容.
22. a. 矩阵的阶梯形式是唯一的.
    b. 矩阵中的主元位置依赖于是否在行化简过程中使用了行互换.
    c. 把矩阵变换为阶梯形式在行化简过程中称为*正向阶段*.
    d. 只要方程组含有自由变量，则解集中含有多个解.
    e. 方程组的通解是方程组所有解的显式表示.
23. 假设一个方程组的 3×5 *系数*矩阵有三个主元列. 该方程组是否相容？为什么？
24. 假设一个线性方程组有 3×5 的*增广*矩阵，该矩阵的第5列为主元列. 该方程组是否相容？为什么？
25. 假设一个线性方程组系数矩阵的每一行中都有一个主元. 说明该方程组为什么相容.
26. 假设一个含三个变量、三个方程的线性方程组，其系数矩阵的每一列都有一个主元. 说明该方程组为什么有唯一解.
27. 用主元列的概念重述定理2的最后一句话："如果一个线性方程组相容，则解唯一当且仅当________."
28. 要想知道线性方程组是相容的，并且有唯一解，我们必须知道哪些关于增广矩阵中主元列的信息？
29. 方程个数比未知量少的一个线性方程组，有时称为一个*亚定组*. 假设这样一个方程组是相容的. 说明为什么亚定组必定有无穷多解.
30. 给出一个含三个未知量、两个方程的相容亚定组的例子.
31. 方程个数比未知量个数多的线性方程组，有时称为*超定组*，这样的方程组是否相容？用一个含两个未知量、三个方程的方程组来说明你的答案.
32. 假设一个 $n\times(n+1)$ 矩阵通过行化简变成简化阶梯形式. 当 $n=30$ 时，反向阶段中包含的操作(浮点运算)占总数的比例大致是多少？若 $n=300$ 呢？

假设实验数据用平面上的点集来表示. 则这些数据的一个**插值多项式**(interpolating polynomial)是指图像经过所有点的一个多项式. 在科学工作中，这样的多项式可以用于估计已知数据点之间的值，也可以用来在计算机屏幕上绘制图形图像的曲线. 求插值多项式的一种方法是求解线性方程组.

26

33. 求数据(1,12)，(2,15)，(3,16)的插值多项式 $p(t)=a_0+a_1t+a_2t^2$. 即求 a_0，a_1 和 a_2，使得

$$a_0+a_1(1)+a_2(1)^2=12$$
$$a_0+a_1(2)+a_2(2)^2=15$$
$$a_0+a_1(3)+a_2(3)^2=16$$

34. [**M**]在风洞试验中，射弹的推动力取决于在不同的速度下测量到的空气阻力：

速度(100英尺/秒)	0	2	4	6	8	10
阻力(100磅)	0	2.90	14.8	39.6	74.3	119

求这些数据的插值多项式，并且估计射弹以750英尺/秒的速度飞行时的推动力. 使用

1. 这类真假判断问题在其他许多章节中都会遇到. 至于如何说明理由，可以参考1.1节习题23和24前面的提示.

$p(t)=a_0+a_1t+a_2t^2+a_3t^3+a_4t^4+a_5t^5$. 如果尝试使用次数小于5的多项式将会出现什么问题？试用3次多项式举例说明.[1]

基础练习答案

1. 增广矩阵的简化阶梯形式和相应的方程组为

$$\begin{bmatrix}1 & 0 & -2 & 9\\ 0 & 1 & 1 & 3\end{bmatrix} \quad \text{和} \quad \begin{aligned} x_1 - 2x_3 &= 9 \\ x_2 + x_3 &= 3 \end{aligned}$$

基本变量为 x_1 和 x_2，通解(见图1-6)为

$$\begin{cases} x_1 = 9 + 2x_3 \\ x_2 = 3 - x_3 \\ x_3 \text{ 是自由变量} \end{cases}$$

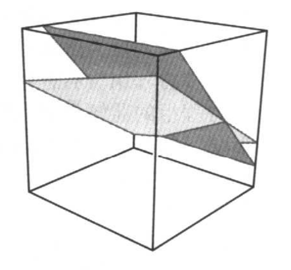


图1-6　线性方程组的通解是两个平面的交线

注意：通解必须表示所有的变量，并且每个参量都要明确标记，这一点非常重要. 下面的陈述并没有表示出解：

$$\begin{cases} x_1 = 9 + 2x_3 \\ x_2 = 3 - x_3 \\ x_3 = 3 - x_2 \end{cases} \quad \text{是错误的解}$$

这个表示意味着 x_2 和 x_3 都是自由的，但实际上不是这样.

2. 行化简方程组的增广矩阵如下：

$$\begin{bmatrix}1 & -2 & -1 & 3 & 0\\ -2 & 4 & 5 & -5 & 3\\ 3 & -6 & -6 & 8 & 2\end{bmatrix} \sim \begin{bmatrix}1 & -2 & -1 & 3 & 0\\ 0 & 0 & 3 & 1 & 3\\ 0 & 0 & -3 & -1 & 2\end{bmatrix} \sim \begin{bmatrix}1 & -2 & -1 & 3 & 0\\ 0 & 0 & 3 & 1 & 3\\ 0 & 0 & 0 & 0 & 5\end{bmatrix}$$

27

这个阶梯矩阵说明方程组是不相容的，因为矩阵最右边的列为主元列；第3行对应于方程 $0=5$. 至此没有必要进行更多的行变换了. 注意方程组是不相容的，所以自由变量的存在并无影响.

1.3　向量方程

线性方程组的重要性质可以用向量的概念和记号来描述. 本节将把含有向量的方程与一般形式的方程组联系起来. 向量这一概念出现在各种数学和物理问题中，我们将在第4章“向量空间”中加以讨论. 在此之前，我们用向量来表示一列数. 这个简单的想法能使我们能够尽快了解向量的一些有趣且重要的应用.

1.3.1　R^2 中的向量

仅含有一列的矩阵称为**列向量**(column vector)，或简称为**向量**(vector). 下面是含两个元素的向量的例子

---

1. 以[M]标记的习题需要在“矩阵软件”辅助下完成，比如MATLAB，Maple，Mathematica，MathCad，Derive，或者一个能处理矩阵的可编程计算器，比如德州仪器公司和惠普公司制造的那一类计算器.

$$\boldsymbol{u}=\begin{bmatrix}3\\-1\end{bmatrix},\quad \boldsymbol{v}=\begin{bmatrix}0.2\\0.3\end{bmatrix},\quad \boldsymbol{w}=\begin{bmatrix}w_1\\w_2\end{bmatrix}$$

其中w_1，w_2为任意实数. 我们把全体含两个元素的向量所构成的集合记为$\mathbf{R}^2$(读作“r－二”). $\mathbf{R}$表示全体实数，它们作为元素在向量中出现. 指数2表示每个向量中含有两个元素.[1]

$\mathbf{R}^2$中的两个向量**相等**(equal)当且仅当向量中对应元素相等. 因此，$\begin{bmatrix}4\\7\end{bmatrix}$不等于$\begin{bmatrix}7\\4\end{bmatrix}$. 我们称$\mathbf{R}^2$中的向量是实数序对.

给定$\boldsymbol{u}$，$\boldsymbol{v}\in\mathbf{R}^2$，将$\boldsymbol{u}$，$\boldsymbol{v}$中对应元素相加，就得到它们的**和**(sum)$\boldsymbol{u}+\boldsymbol{v}$. 例如

$$\begin{bmatrix}1\\-2\end{bmatrix}+\begin{bmatrix}2\\5\end{bmatrix}=\begin{bmatrix}1+2\\-2+5\end{bmatrix}=\begin{bmatrix}3\\3\end{bmatrix}$$

给定向量$\boldsymbol{u}$和实数c，将$\boldsymbol{u}$中每个元素同乘以c，得到c乘$\boldsymbol{u}$的**数量倍数**(scalar multiple，也称作数乘). 例如，

$$若\ \boldsymbol{u}=\begin{bmatrix}3\\-1\end{bmatrix}且\ c=5，则\ c\boldsymbol{u}=5\begin{bmatrix}3\\-1\end{bmatrix}=\begin{bmatrix}15\\-5\end{bmatrix}$$

28

$c\boldsymbol{u}$中的数c称为**数量**(scalar)，为了区别于粗体字$\boldsymbol{u}$，c用细体字表示.

如下面例题所示，数乘可以与向量加法组合运算.

【例题1】 已知$\boldsymbol{u}=\begin{bmatrix}1\\-2\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}2\\-5\end{bmatrix}$，求$4\boldsymbol{u}$，$(-3)\boldsymbol{v}$和$4\boldsymbol{u}+(-3)\boldsymbol{v}$.

解：

$$4\boldsymbol{u}=\begin{bmatrix}4\\-8\end{bmatrix},\quad (-3)\boldsymbol{v}=\begin{bmatrix}-6\\15\end{bmatrix}$$

且

$$4\boldsymbol{u}+(-3)\boldsymbol{v}=\begin{bmatrix}4\\-8\end{bmatrix}+\begin{bmatrix}-6\\15\end{bmatrix}=\begin{bmatrix}-2\\7\end{bmatrix}$$ ■

为了方便(同时也为了节省篇幅)，我们有时将列向量$\begin{bmatrix}3\\-1\end{bmatrix}$写成$(3,-1)$的形式. 这时，我们用圆括号和逗号来表示该列向量，以区别于1×2行矩阵$[3\ \ -1]$. 尽管下面两个矩阵包含相同的元素，但由于形式不同，因此

1. 本书大部分内容中的向量和矩阵仅含有实元素. 然而，第1章到第5章的所有概念和定理，以及其余的大部分内容在元素为复数时仍然成立. 例如，在电机工程和物理学中会很自然地出现复数向量和复数矩阵.

$$\begin{bmatrix} 3 \\ -1 \end{bmatrix} \neq [\,3 \;\; -1\,]$$

1.3.2　$\mathbf{R}^2$的几何表示

考虑平面直角坐标系．由于平面中的每一点都由一个有序数对确定，我们就可以把几何点(a,b)看成列向量$\begin{bmatrix} a \\ b \end{bmatrix}$．因此可以把$\mathbf{R}^2$看作平面上所有点的集合．见图1-7．

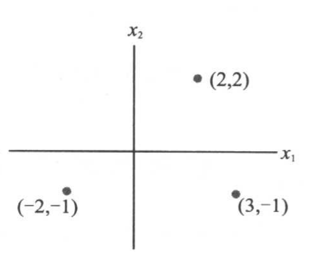


图 1-7　将向量看作点

如图 1-8 所示，向量$\begin{bmatrix} 3 \\ -1 \end{bmatrix}$的几何直观就是一个从原点$(0,0)$指向点$(3,-1)$的箭头（有向线段）．这种情况下，单点没有特别意义．[1]

两个向量之和在几何表示中很有意义，下列法则可用解析几何加以验证．

29

加法的平行四边形法则

如果$\boldsymbol{u}$，$\boldsymbol{v} \in \mathbf{R}^2$表示平面上的点，则$\boldsymbol{u}+\boldsymbol{v}$是$\mathbf{0}$，$\boldsymbol{u}$，$\boldsymbol{v}$所确定的平行四边形的第四个顶点．见图 1-9．

【例题 2】　向量$\boldsymbol{u}=\begin{bmatrix} 2 \\ 2 \end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix} -6 \\ 1 \end{bmatrix}$及$\boldsymbol{u}+\boldsymbol{v}=\begin{bmatrix} -4 \\ 3 \end{bmatrix}$，如图 1-10 所示．■

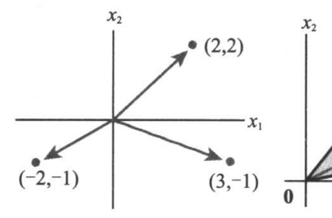


图 1-8　带箭头的向量

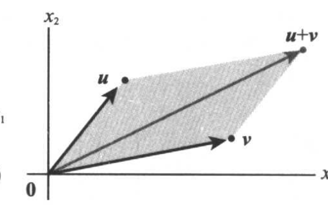


图 1-9　平行四边形法则

图　1-10

下面的例题表明，非零向量的全体数量倍数构成一条过原点$(0,0)$的直线．

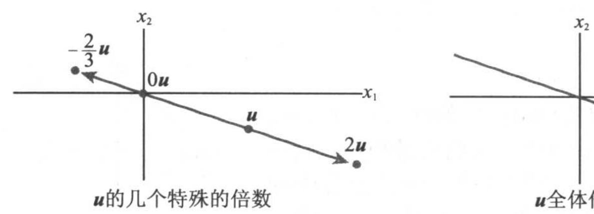
【例题 3】　设$\boldsymbol{u}=\begin{bmatrix} 3 \\ -1 \end{bmatrix}$．在图上画出向量$\boldsymbol{u}$，$2\boldsymbol{u}$和$-\frac{2}{3}\boldsymbol{u}$.

解：图 1-11 已经标出$\boldsymbol{u}$，$2\boldsymbol{u}=\begin{bmatrix} 6 \\ -2 \end{bmatrix}$及$-\frac{2}{3}\boldsymbol{u}=\begin{bmatrix} -2 \\ 2/3 \end{bmatrix}$．箭头$2\boldsymbol{u}$的长度是箭头$\boldsymbol{u}$的两倍，且两者同向．箭头$-\frac{2}{3}\boldsymbol{u}$的长度是箭头$\boldsymbol{u}$的2/3，且两者反向．一般来说，箭头

1. 在物理学中，箭头可以表示力，并且可以在空间中自由移动．向量的这种解释将在 4.1 节中讨论．

$c\boldsymbol{u}$ 的长度是箭头 $\boldsymbol{u}$ 的 $|c|$ 倍. [回忆 $(0,0)$ 到 (a,b) 的线段长度为 $\sqrt{a^2+b^2}$，这个问题在第 6 章中将会进一步讨论.] ■

30



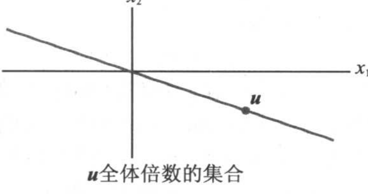


图 1-11

1.3.3 $\mathbf{R}^3$中的向量

$\mathbf{R}^3$中的向量是含有三个元素的 3×1 矩阵. 在几何上，它们可以直观地表示为三维坐标空间的点，为使这种表示更清晰，还可配以从原点出发到这些点的箭头. 图 1-12 中绘出了向量 $\boldsymbol{a}=\begin{bmatrix}2\\3\\4\end{bmatrix}$ 和 $2\boldsymbol{a}$.

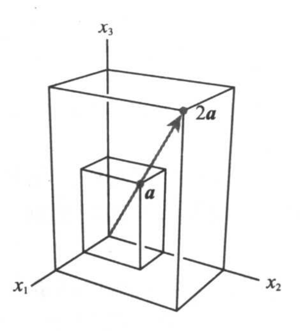


图 1-12 $\mathbf{R}^3$ 中的数乘

1.3.4 $\mathbf{R}^n$中的向量

若 n 为正整数，$\mathbf{R}^n$（读作“r - n”）表示全体 n 元实数列（或有序 n 元组）构成的集合. $\mathbf{R}^n$中的向量通常写成如下形式的 $n\times1$ 列矩阵：

$$\boldsymbol{u}=\begin{bmatrix}u_1\\u_2\\\vdots\\u_n\end{bmatrix}$$

元素全为零的向量称为**零向量**（zero vector），记为 $\mathbf{0}$（向量 $\mathbf{0}$ 中元素的个数可根据上下文得出）.

与 $\mathbf{R}^2$中一样，$\mathbf{R}^n$中向量相等、数乘及加法运算也都根据其元素来定义. 向量的这些运算满足以下性质，这些性质可以用实数的性质直接验证. 见本节末尾基础练习 1 以及习题 33 和 34.

31

$\mathbf{R}^n$的代数性质

对 $\mathbf{R}^n$中任意向量 $\boldsymbol{u}$，$\boldsymbol{v}$，$\boldsymbol{w}$ 和任意数量 c，d，都有：

(i) $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$	(v) $c(\boldsymbol{u}+\boldsymbol{v})=c\boldsymbol{u}+c\boldsymbol{v}$
(ii) $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$	(vi) $(c+d)\boldsymbol{u}=c\boldsymbol{u}+d\boldsymbol{u}$
(iii) $\boldsymbol{u}+\mathbf{0}=\mathbf{0}+\boldsymbol{u}=\boldsymbol{u}$	(vii) $c(d\boldsymbol{u})=(cd)(\boldsymbol{u})$
(iv) $\boldsymbol{u}+(-\boldsymbol{u})=-\boldsymbol{u}+\boldsymbol{u}=\mathbf{0}$	(viii) $1\boldsymbol{u}=\boldsymbol{u}$

其中 $-\boldsymbol{u}$ 表示 $(-1)\boldsymbol{u}$

为了简化记号，我们使用“向量减法”，即用 $\boldsymbol{u}-\boldsymbol{v}$ 代替 $\boldsymbol{u}+(-1)\boldsymbol{v}$. 图 1-13 说

明 $\boldsymbol{u}-\boldsymbol{v}$ 等同于 $\boldsymbol{u}+(-1)\boldsymbol{v}$.

1.3.5　线性组合

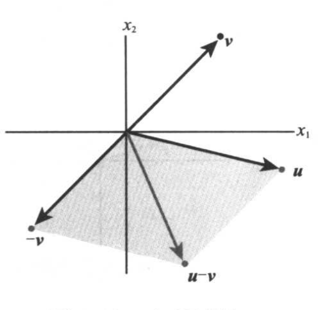


图 1-13　向量减法

给定 $\mathbf{R}^n$ 中向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p$，以及数量 $c_1,c_2,\cdots,c_p$，定义向量 $\boldsymbol{y}$ 为：

$$\boldsymbol{y}=c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p$$

则 $\boldsymbol{y}$ 称为以 $c_1,c_2,\cdots,c_p$ 为**权重**(weight)的向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p$ 的**线性组合**(linear combination). 前面的性质(ii)允许我们在构造这个线性组合时省略圆括号. 线性组合中的权重可以是任意实数，包括0. 例如，向量 $\boldsymbol{v}_1$，$\boldsymbol{v}_2$ 的一些线性组合为

$$\sqrt{3}\boldsymbol{v}_1+\boldsymbol{v}_2,\quad \frac{1}{2}\boldsymbol{v}_1\left(=\frac{1}{2}\boldsymbol{v}_1+0\boldsymbol{v}_2\right),\quad \boldsymbol{0}(=0\boldsymbol{v}_1+0\boldsymbol{v}_2)$$

【例题4】　图1-14给出了向量 $\boldsymbol{v}_1=\begin{bmatrix}-1\\1\end{bmatrix}$ 和 $\boldsymbol{v}_2=\begin{bmatrix}2\\1\end{bmatrix}$ 的若干线性组合.（注意过 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 整数倍的平行网格.）估计生成向量 $\boldsymbol{u}$ 和 $\boldsymbol{w}$ 的 $\boldsymbol{v}_1$，$\boldsymbol{v}_2$ 的线性组合.

32

解： 根据平行四边形法则，$\boldsymbol{u}$ 是 $3\boldsymbol{v}_1$ 和 $-2\boldsymbol{v}_2$ 的和，即

$$\boldsymbol{u}=3\boldsymbol{v}_1-2\boldsymbol{v}_2$$

$\boldsymbol{u}$ 的线性组合表示可以解释为从原点出发沿着直线路径行进的指令. 首先，向 $\boldsymbol{v}_1$ 的方向行进3个单位到达 $3\boldsymbol{v}_1$，然后向 $\boldsymbol{v}_2$ 的方向(平行于过0点和 $\boldsymbol{v}_2$ 点的直线方向)行进 -2 个单位. 下一步，虽然向量 $\boldsymbol{w}$ 不在网格线上，却大概位于两对平行线的正中间，是 $(5/2)\boldsymbol{v}_1$ 和 $(-1/2)\boldsymbol{v}_2$ 确定的平行四边形的一个顶点(见图1-15). 因此

$$\boldsymbol{w}=\frac{5}{2}\boldsymbol{v}_1-\frac{1}{2}\boldsymbol{v}_2$$

■

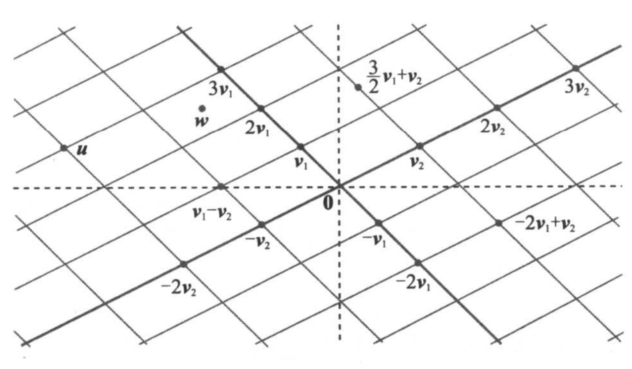


图1-14　$\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的线性组合

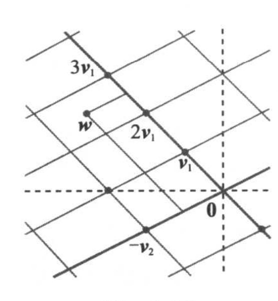


图　1-15

下面的例题将线性组合与1.1，1.2节中的解的存在性问题联系起来.

【例题5】　设 $\boldsymbol{a}_1=\begin{bmatrix}1\\-2\\-5\end{bmatrix}$，$\boldsymbol{a}_2=\begin{bmatrix}2\\5\\6\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}7\\4\\-3\end{bmatrix}$. 判断 $\boldsymbol{b}$ 是否能由 $\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2$ 的线性组合

生成(表示)．即，判断是否存在 x_1 和 x_2 使得

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 = \boldsymbol{b} \tag{1}$$

如果向量方程(1)有解，求其解.

解：用数乘和向量加法的定义来重写向量方程为

$$x_1\begin{bmatrix}1\\-2\\-5\end{bmatrix}+x_2\begin{bmatrix}2\\5\\6\end{bmatrix}=\begin{bmatrix}7\\4\\-3\end{bmatrix}$$

$$\uparrow\ \boldsymbol{a}_1 \qquad \uparrow\ \boldsymbol{a}_2 \qquad \uparrow\ \boldsymbol{b}$$

等同于

$$\begin{bmatrix}x_1\\-2x_1\\-5x_1\end{bmatrix}+\begin{bmatrix}2x_2\\5x_2\\6x_2\end{bmatrix}=\begin{bmatrix}7\\4\\-3\end{bmatrix}$$

和

$$\begin{bmatrix}x_1+2x_2\\-2x_1+5x_2\\-5x_1+6x_2\end{bmatrix}=\begin{bmatrix}7\\4\\-3\end{bmatrix} \tag{2}$$

方程(2)中左右两端的向量相等当且仅当对应的元素都分别相等．即，x_1 和 x_2 使得方程(1)成立当且仅当 x_1 和 x_2 满足方程组

33

$$\begin{aligned} x_1+2x_2 &= 7\\ -2x_1+5x_2 &= 4\\ -5x_1+6x_2 &= -3 \end{aligned} \tag{3}$$

对方程组的增广矩阵进行下列行化简，以求解该方程组：[1]

$$\begin{bmatrix}1&2&7\\-2&5&4\\-5&6&-3\end{bmatrix}\sim\begin{bmatrix}1&2&7\\0&9&18\\0&16&32\end{bmatrix}\sim\begin{bmatrix}1&2&7\\0&1&2\\0&16&32\end{bmatrix}\sim\begin{bmatrix}1&0&3\\0&1&2\\0&0&0\end{bmatrix}$$

方程组(3)的解是 $x_1=3$ 和 $x_2=2$．因此 $\boldsymbol{b}$ 是 $\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2$ 的线性组合，权重为 $x_1=3$ 和 $x_2=2$，即，

$$3\begin{bmatrix}1\\-2\\-5\end{bmatrix}+2\begin{bmatrix}2\\5\\6\end{bmatrix}=\begin{bmatrix}7\\4\\-3\end{bmatrix}$$ ■

注意，例题5中的初始向量 $\boldsymbol{a}_1$，$\boldsymbol{a}_2$ 和 $\boldsymbol{b}$ 就是被我们施以行化简的增广矩阵的列：

$$\begin{bmatrix}1&2&7\\-2&5&4\\-5&6&-3\end{bmatrix}$$

$$\uparrow\ \boldsymbol{a}_1 \quad \uparrow\ \boldsymbol{a}_2 \quad \uparrow\ \boldsymbol{b}$$

1. 矩阵之间的符号 ~ 表示行等价(1.2节).

为强调矩阵中的列，我们可将该矩阵写作：

$$[\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \boldsymbol{b}] \tag{4}$$

很明显，不需要经过例题5的中间步骤，可以根据向量方程(1)立即写出增广矩阵. 如(4)所示，只需按照向量在(1)中出现的次序排列成矩阵的列即可.

上述讨论经过简单修改后可得到以下重要的结论.

向量方程

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{b}$$

与增广矩阵为

$$[\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \cdots \quad \boldsymbol{a}_n \quad \boldsymbol{b}] \tag{5}$$

的线性方程组有相同的解集. 特别地，$\boldsymbol{b}$ 可以由 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$ 的线性组合表示当且仅当(5)对应的线性方程组有解.

34

线性代数的一个关键思想就是研究由指定向量集$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$的全体线性组合所构成的集合.

【定义】　如果$\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p \in \mathbf{R}^n$，则所有由$\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p$生成的全体线性组合的集合记为$\mathrm{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$，称作由$\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p$张成(或生成)的$\mathbf{R}^n$的子空间. 即，$\mathrm{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$是所有可以写成如下形式的向量构成的集合：

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_p\boldsymbol{v}_p$$

其中，$c_1,\cdots,c_p$为数量.

向量 $\boldsymbol{b}$ 是否属于 $\mathrm{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p\}$等同于向量方程

$$x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_p\boldsymbol{v}_p = \boldsymbol{b}$$

是否有解，或者增广矩阵$[\boldsymbol{v}_1\cdots\boldsymbol{v}_p\boldsymbol{b}]$对应的线性方程组是否有解.

注意，$\mathrm{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$包含了$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$中向量的全体数量倍数，因为以 $\boldsymbol{v}_1$ 为例，我们有 $c\boldsymbol{v}_1 = c\boldsymbol{v}_1 + 0\boldsymbol{v}_2 + \cdots + 0\boldsymbol{v}_p$. 特别地，零向量一定属于 $\mathrm{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$.

1.3.6　Span{v}和 Span{u, v}的几何表示

设 $\boldsymbol{v}$ 是 $\mathbf{R}^3$ 中的一个非零向量. 则 $\mathrm{Span}\{\boldsymbol{v}\}$是 $\boldsymbol{v}$ 的全体数量倍数构成的集合，我们将 $\mathrm{Span}\{\boldsymbol{v}\}$直观地看成 $\mathbf{R}^3$ 中过 $\mathbf{0}$ 和 $\boldsymbol{v}$ 的直线上的点集. 见图1-16.

如果 $\boldsymbol{u}$，$\boldsymbol{v}$ 是 $\mathbf{R}^3$ 中的非零向量，$\boldsymbol{v}$ 不是 $\boldsymbol{u}$ 的倍数，则 $\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$是 $\mathbf{R}^3$ 中包含 $\boldsymbol{u}$，$\boldsymbol{v}$，$\mathbf{0}$ 的平面. 特别地，$\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$包含 $\mathbf{R}^3$ 中分别过 $\boldsymbol{u}$，$\mathbf{0}$ 以及 $\boldsymbol{v}$，$\mathbf{0}$ 的两条直线. 见图1-17.

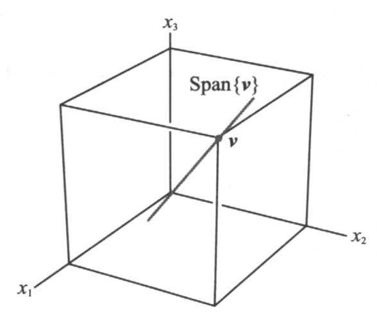


图1-16　$\mathrm{Span}\{\boldsymbol{v}\}$是过原点的直线

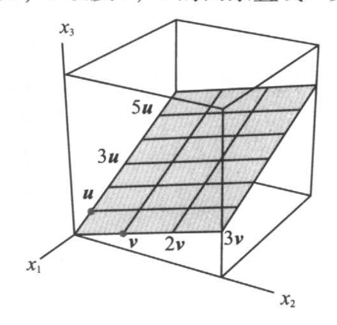


图1-17　$\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$是过原点的平面

35

【例题6】 设 $\boldsymbol{a}_1=\begin{bmatrix}1\\-2\\3\end{bmatrix}$, $\boldsymbol{a}_2=\begin{bmatrix}5\\-13\\-3\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}-3\\8\\1\end{bmatrix}$. 则 Span$\{\boldsymbol{a}_1,\boldsymbol{a}_2\}$为 $\mathbf{R}^3$ 中过原点的平面. $\boldsymbol{b}$ 是否属于该平面?

解: 方程 $x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2=\boldsymbol{b}$ 是否有解? 为了回答这个问题, 对矩阵$[\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \boldsymbol{b}]$进行行化简:

$$\begin{bmatrix}1&5&-3\\-2&-13&8\\3&-3&1\end{bmatrix}\sim\begin{bmatrix}1&5&-3\\0&-3&2\\0&-18&10\end{bmatrix}\sim\begin{bmatrix}1&5&-3\\0&-3&2\\0&0&-2\end{bmatrix}$$

第三个方程为 $0x_2=-2$, 说明方程组无解. 因此向量方程 $x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2=\boldsymbol{b}$ 无解, 即 $\boldsymbol{b}$ 不属于 Span$\{\boldsymbol{a}_1,\boldsymbol{a}_2\}$. ■

1.3.7 线性组合的应用

最后一道例题说明, 当我们将某个量(例如"成本")分解成几部分时要如何用到数量倍数和线性组合. 例题的基本前提是已知单位成本, 求总成本:

$$\{单位数量\}\cdot\{单位成本\}=\{总成本\}$$

【例题7】 某公司生产两种产品. 生产价值 1.00 美元的产品 B 需花费原材料 0.45 美元, 劳动力 0.25 美元, 管理费 0.15 美元. 生产价值 1.00 美元的产品 C 需花费原材料 0.40 美元, 劳动力 0.30 美元, 管理费 0.15 美元. 设

$$\boldsymbol{b}=\begin{bmatrix}0.45\\0.25\\0.15\end{bmatrix}和\ \boldsymbol{c}=\begin{bmatrix}0.40\\0.30\\0.15\end{bmatrix}$$

则 $\boldsymbol{b}$ 和 $\boldsymbol{c}$ 分别代表两种产品的"一美元产出成本".

a. 向量 $100\boldsymbol{b}$ 的经济学解释是什么?

b. 假设该公司希望生产价值 x_1美元的产品 B 和价值 x_2美元的产品 C. 请写出表示该公司需要花费的各种成本(原材料、劳动力、管理费)的向量.

解:

a. 我们有

$$100\boldsymbol{b}=100\begin{bmatrix}0.45\\0.25\\0.15\end{bmatrix}=\begin{bmatrix}45\\25\\15\end{bmatrix}$$

向量 $100\boldsymbol{b}$ 表示了生产价值 100 美元产品 B 的不同成本花费, 即, 原材料 45 美元, 劳动力 25 美元, 管理费 15 美元.

36

b. 生产价值 x_1美元的产品 B, 成本用 $x_1\boldsymbol{b}$ 表示, 生产价值 x_2美元的产品 C, 成本用 $x_2\boldsymbol{c}$ 表示. 因此生产这两种产品的总成本为 $x_1\boldsymbol{b}+x_2\boldsymbol{c}$. ■

基础练习

1. 证明对任意 $\boldsymbol{u}$, $\boldsymbol{v}\in\mathbf{R}^n$, $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$ 成立.
2. h 取何值, $\boldsymbol{y}$ 属于 Span$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$. 其中

$$v_1=\begin{bmatrix}1\\-1\\-2\end{bmatrix},\quad v_2=\begin{bmatrix}5\\-4\\-7\end{bmatrix},\quad v_3=\begin{bmatrix}-3\\1\\0\end{bmatrix},\quad y=\begin{bmatrix}-4\\3\\h\end{bmatrix}$$

习题 1.3

在习题1和2中，计算 $\boldsymbol{u}+\boldsymbol{v}$ 和 $\boldsymbol{u}-2\boldsymbol{v}$.

1. $\boldsymbol{u}=\begin{bmatrix}-1\\2\end{bmatrix},\quad \boldsymbol{v}=\begin{bmatrix}-3\\-1\end{bmatrix}$

2. $\boldsymbol{u}=\begin{bmatrix}3\\2\end{bmatrix},\quad \boldsymbol{v}=\begin{bmatrix}2\\-1\end{bmatrix}$

习题3和4，要求在 xy-坐标中标出下列向量：$\boldsymbol{u}$，$\boldsymbol{v}$，$-\boldsymbol{v}$，$-2\boldsymbol{v}$，$\boldsymbol{u}+\boldsymbol{v}$，$\boldsymbol{u}-\boldsymbol{v}$ 和 $\boldsymbol{u}-2\boldsymbol{v}$. 注意 $\boldsymbol{u}-\boldsymbol{v}$ 是平行四边形的一个顶点，该平行四边形的其余顶点是 $\boldsymbol{u}$，$\boldsymbol{0}$，$-\boldsymbol{v}$.

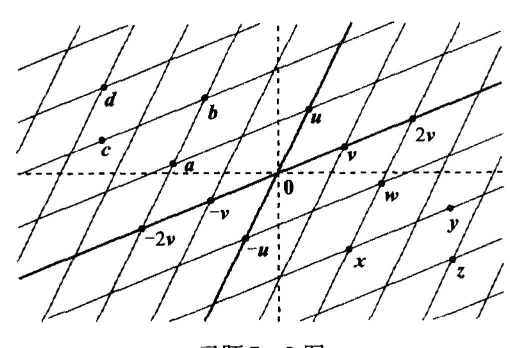
3. 习题1中的 $\boldsymbol{u}$ 和 $\boldsymbol{v}$.
4. 习题2中的 $\boldsymbol{u}$ 和 $\boldsymbol{v}$.

在习题5和6中，写出一个与已知向量方程等价的方程组.

5. $x_1\begin{bmatrix}6\\-1\\5\end{bmatrix}+x_2\begin{bmatrix}-3\\4\\0\end{bmatrix}=\begin{bmatrix}1\\-7\\-5\end{bmatrix}$

6. $x_1\begin{bmatrix}-2\\3\end{bmatrix}+x_2\begin{bmatrix}8\\5\end{bmatrix}+x_3\begin{bmatrix}1\\-6\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$

根据下图，把习题7和8中列出的向量写成 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的线性组合形式. $\mathbf{R}^2$ 中的每个向量是否都是 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的线性组合？



习题7、8图

7. 向量 $\boldsymbol{a}$，$\boldsymbol{b}$，$\boldsymbol{c}$ 和 $\boldsymbol{d}$
8. 向量 $\boldsymbol{w}$，$\boldsymbol{x}$，$\boldsymbol{y}$ 和 $\boldsymbol{z}$

在习题9和10中，写出与已知方程组等价的向量方程.

9.
$$\begin{aligned} x_2+5x_3&=0\\ 4x_1+6x_2-x_3&=0\\ -x_1+3x_2-8x_3&=0\end{aligned}$$

10.
$$\begin{aligned} 4x_1+x_2+3x_3&=9\\ x_1-7x_2-2x_3&=2\\ 8x_1+6x_2-5x_3&=15\end{aligned}$$

[37]

在习题11和12中，判断 $\boldsymbol{b}$ 是否为 $\boldsymbol{a}_1$，$\boldsymbol{a}_2$ 和 $\boldsymbol{a}_3$ 的线性组合.

11. $\boldsymbol{a}_1=\begin{bmatrix}1\\-2\\0\end{bmatrix},\quad \boldsymbol{a}_2=\begin{bmatrix}0\\1\\2\end{bmatrix},\quad \boldsymbol{a}_3=\begin{bmatrix}5\\-6\\8\end{bmatrix},\quad \boldsymbol{b}=\begin{bmatrix}2\\-1\\6\end{bmatrix}$

12. $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$, $\boldsymbol{a}_2 = \begin{bmatrix} 0 \\ 5 \\ 5 \end{bmatrix}$, $\boldsymbol{a}_3 = \begin{bmatrix} 2 \\ 0 \\ 8 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} -5 \\ 11 \\ -7 \end{bmatrix}$

在习题 13 和 14 中，判断 $\boldsymbol{b}$ 是否是矩阵 $\boldsymbol{A}$ 的列向量的线性组合.

13. $\boldsymbol{A} = \begin{bmatrix} 1 & -4 & 2 \\ 0 & 3 & 5 \\ -2 & 8 & -4 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 3 \\ -7 \\ -3 \end{bmatrix}$

14. $\boldsymbol{A} = \begin{bmatrix} 1 & -2 & -6 \\ 0 & 3 & 7 \\ 1 & -2 & 5 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 11 \\ -5 \\ 9 \end{bmatrix}$

在习题 15 和 16 中，请列出 Span$\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$中的 5 个向量. 说明生成每一向量的 $\boldsymbol{v}_1$，$\boldsymbol{v}_2$ 的权重，并且写出向量的三个元素. 不要画图.

15. $\boldsymbol{v}_1 = \begin{bmatrix} 7 \\ 1 \\ -6 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} -5 \\ 3 \\ 0 \end{bmatrix}$

16. $\boldsymbol{v}_1 = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} -2 \\ 0 \\ 3 \end{bmatrix}$

17. 设 $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix}$, $\boldsymbol{a}_2 = \begin{bmatrix} -2 \\ -3 \\ 7 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 4 \\ 1 \\ h \end{bmatrix}$. h 取何值时，$\boldsymbol{b}$ 在 $\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2$ 生成的子空间中？

18. 设 $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} -3 \\ 1 \\ 8 \end{bmatrix}$, $\boldsymbol{y} = \begin{bmatrix} h \\ -5 \\ -3 \end{bmatrix}$. h 取何值时，$\boldsymbol{y}$ 在 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 生成的子空间中？

19. 设 $\boldsymbol{v}_1 = \begin{bmatrix} 8 \\ 2 \\ -6 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} 12 \\ 3 \\ -9 \end{bmatrix}$, 给出 Span$\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$的几何表示.

20. 对习题 16 中的向量给出 Span$\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$的几何表示.

21. 设 $\boldsymbol{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\boldsymbol{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. 证明 h，k 取任意值，向量$\begin{bmatrix} h \\ k \end{bmatrix}$都在 Span$\{\boldsymbol{u}, \boldsymbol{v}\}$中.

22. 构造一个没有零元素的 3×3 矩阵 $\boldsymbol{A}$，以及 $\mathbf{R}^3$ 中的一个向量 $\boldsymbol{b}$，该向量不在由 $\boldsymbol{A}$ 的列生成的子空间中.

在习题 23 和 24 中，判断每个命题的真假，并说明理由.

23. a. 向量$\begin{bmatrix} -4 \\ 3 \end{bmatrix}$的另一个表示是$[-4 \quad 3]$.

b. 平面中对应于$\begin{bmatrix} -2 \\ 5 \end{bmatrix}$和$\begin{bmatrix} -5 \\ 2 \end{bmatrix}$的点在过原点的一条直线上.

c. $\frac{1}{2}\boldsymbol{v}_1$是 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的一个线性组合.

d. 增广矩阵$[\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \boldsymbol{a}_3 \quad \boldsymbol{b}]$对应的线性方程组与方程 $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3 = \boldsymbol{b}$ 的解集相同.

e. 集合 Span$\{\boldsymbol{u}, \boldsymbol{v}\}$总可以形象化为一个过原点的平面.

24. a. 任意 5 个实数构成的序列都是 $\mathbf{R}^5$ 中的一个向量.

b. 向量 $\boldsymbol{u} - \boldsymbol{v}$ 与向量 $\boldsymbol{v}$ 相加得到向量 $\boldsymbol{u}$.

c. 线性组合 $c_1\boldsymbol{v}_1 + \cdots + c_p\boldsymbol{v}_p$中的权重 $c_1, \cdots, c_p$不能全为零.

d. 当 $\boldsymbol{u}$，$\boldsymbol{v}$ 为非零向量时，Span$\{\boldsymbol{u}, \boldsymbol{v}\}$包含过原点和 $\boldsymbol{u}$ 的直线.

e. 增广矩阵为$[\boldsymbol{a}_1\quad \boldsymbol{a}_2\quad \boldsymbol{a}_3\quad \boldsymbol{b}]$的线性方程组是否有解等同于$\boldsymbol{b}$是否在$\text{Span}\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3\}$中.

25. 设$\boldsymbol{A}=\begin{bmatrix}1&0&-4\\0&3&-2\\-2&6&3\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}4\\1\\-4\end{bmatrix}$. 用$\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$表示A中的列，并设$W=\text{Span}\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3\}$.

a. $\boldsymbol{b}$是否在$\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3\}$中？ $\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3\}$中有多少个向量？

b. $\boldsymbol{b}$是否在$\boldsymbol{W}$中？$\boldsymbol{W}$中有多少个向量？

c. 证明$\boldsymbol{a}_1$在$\boldsymbol{W}$中. [提示：不用进行行变换.]

26. 设$\boldsymbol{A}=\begin{bmatrix}2&0&6\\-1&8&5\\1&-2&1\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}10\\3\\3\end{bmatrix}$. 并且设$\boldsymbol{W}$为$\boldsymbol{A}$中列的所有线性组合构成的集合.

a. $\boldsymbol{b}$是否在$\boldsymbol{W}$中？

b. 说明$\boldsymbol{A}$中的第3列在$\boldsymbol{W}$中.

27. 一个采矿公司有两个矿井，矿井#1每天开采的矿产中含20吨铜矿和550公斤银矿，矿井#2每天开采的矿产中含有30吨铜矿和500公斤银矿. 设$\boldsymbol{v}_1=\begin{bmatrix}20\\550\end{bmatrix}$, $\boldsymbol{v}_2=\begin{bmatrix}30\\500\end{bmatrix}$. 则$\boldsymbol{v}_1$和$\boldsymbol{v}_2$分别表示矿井#1和矿井#2“每天的产量”.

38

a. 向量$5\boldsymbol{v}_1$的实际意义是什么？

b. 假设公司对矿井#1和矿井#2分别开采$\boldsymbol{x}_1$天和$\boldsymbol{x}_2$天. 写出向量方程，使方程的解表示开采150吨铜矿和2825公斤银矿时，每个矿井需要开采的天数. 不用解方程.

c. [**M**]求解(b)中的方程.

28. 一个蒸汽发电厂用以下两种煤作燃料：无烟煤(A)和烟煤(B). 每燃烧一吨A，发电厂能产出2.76千万比特热量、3100克二氧化硫和250克微粒物质(固体微粒污染物). 每燃烧一吨B，发电厂能产出3.02千万比特热量，6400克二氧化硫和360克微粒物质.

a. 蒸汽发电厂燃烧x_1吨A和x_2吨B能产出多少热量？

b. 若蒸汽发电厂的产量用比特热量、二氧化硫和微粒物质的向量来表示，假设燃烧x_1吨的A和x_2吨的B，请把产量表示成两个向量的线性组合形式.

c. [**M**]在某一个时段里，蒸汽发电厂产出16.2千万比特热量，23 610克二氧化硫和1623克微粒物质. 判断每一种燃料发电厂分别燃烧了多少吨. 向量方程应作为解答中的一部分.

29. 设$\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k$是$\mathbf{R}^3$中的点，假设对于$j=1,\cdots,k$，质量为$m_j$的物体位于$\boldsymbol{v}_j$点上. 物理上称为物体的**点质量**. 点质量系统的总质量为

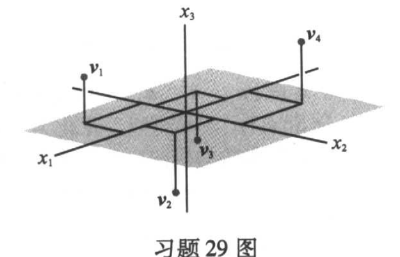
$$m=m_1+\cdots+m_k$$

这个系统的**重心**(或**质心**)为

$$\bar{\boldsymbol{v}}=\frac{1}{m}[m_1\boldsymbol{v}_1+\cdots+m_k\boldsymbol{v}_k]$$

计算包含下列点质量系统的中心(见下图)：

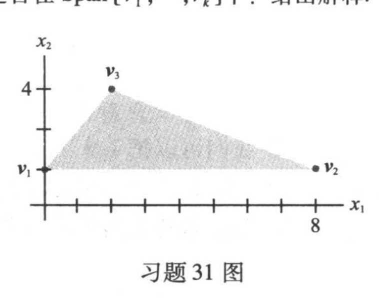
点	质量
$\boldsymbol{v}_1=(5,-4,3)$	2g
$\boldsymbol{v}_2=(4,3,-2)$	5g
$\boldsymbol{v}_3=(-4,-3,-1)$	2g
$\boldsymbol{v}_4=(-9,8,6)$	1g



习题29图

30. 设 $\boldsymbol{v}$ 是习题 29 中 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k$的点质量系统的重心. $\boldsymbol{v}$ 是否在 Span$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k\}$中？给出解释.

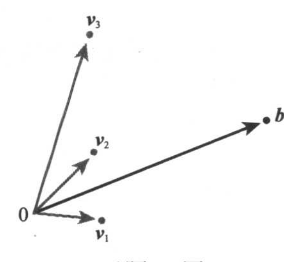
31. 一块密度均匀、厚度均匀的薄三角板顶点为 $\boldsymbol{v}_1=(0,1)$，$\boldsymbol{v}_2=(8,1)$和 $\boldsymbol{v}_3=(2,4)$，如右图所示，且三角板的质量为 3 克.
   a. 找出三角板质量中心(x,y)的坐标. 该"平衡点"与三角板顶点的三个 1 - 克点质量组成的质量系统的中心一致.
   b. 如何把 6 克的质量分配到三角板的顶点上，使得三角板的平衡点移动到$(2,2)$上？[提示：设 w_1，w_2，w_3分别表示加到三个顶点的质量，有 $w_1+w_2+w_3=6$.]



习题 31 图

39

32. 考虑 $\mathbf{R}^2$ 中的向量 $\boldsymbol{v}_1$，$\boldsymbol{v}_2$，$\boldsymbol{v}_3$和 $\boldsymbol{b}$，如右图所示. 方程 $x_1\boldsymbol{v}_1+x_2\boldsymbol{v}_2+x_3\boldsymbol{v}_3=\boldsymbol{b}$ 是否有解？解是否唯一？利用图示解释你的答案.



习题 32 图

33. 用向量 $\boldsymbol{u}=(u_1,\cdots,u_n)$，$\boldsymbol{v}=(v_1,\cdots,v_n)$和 $\boldsymbol{w}=(w_1,\cdots,w_n)$来验证 $\mathbf{R}^n$的下列代数性质.
   a. $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$
   b. $c(\boldsymbol{u}+\boldsymbol{v})=c\boldsymbol{u}+c\boldsymbol{v}$，　c 为任意数量

34. 用向量 $\boldsymbol{u}=(u_1,\cdots,u_n)$验证 $\mathbf{R}^n$中的如下代数性质.
   a. $\boldsymbol{u}+(-\boldsymbol{u})=(-\boldsymbol{u})+\boldsymbol{u}=\mathbf{0}$
   b. $c(d\boldsymbol{u})=(cd)\boldsymbol{u}$，　c 和 d 为任意数量.

基础练习答案

1. 取 $\mathbf{R}^n$中任意向量 $\boldsymbol{u}=(u_1,\cdots,u_n)$和 $\boldsymbol{v}=(v_1,\cdots,v_n)$，计算

$$
\begin{aligned}
\boldsymbol{u}+\boldsymbol{v} &=(u_1+v_1,\cdots,u_n+v_n) && \text{向量加法的定义}\\
&=(v_1+u_1,\cdots,v_n+u_n) && \mathbf{R}\text{ 中加法交换律}\\
&=\boldsymbol{v}+\boldsymbol{u} && \text{向量加法的定义}
\end{aligned}
$$

2. 向量 $\boldsymbol{y}$ 属于 Span$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$当且仅当存在 x_1，x_2，x_3使得

$$
x_1\begin{bmatrix}1\\-1\\-2\end{bmatrix}+x_2\begin{bmatrix}5\\-4\\-7\end{bmatrix}+x_3\begin{bmatrix}-3\\1\\0\end{bmatrix}=\begin{bmatrix}-4\\3\\h\end{bmatrix}
$$

这个方程与一个含三个未知量的三个线性方程组成的方程组等价. 如果对这个方程组的增广矩阵进行行化简，将发现

$$
\begin{bmatrix}1&5&-3&-4\\-1&-4&1&3\\-2&-7&0&h\end{bmatrix}\sim\begin{bmatrix}1&5&-3&-4\\0&1&-2&-1\\0&3&-6&h-8\end{bmatrix}
$$

$$
\sim\begin{bmatrix}1&5&-3&-4\\0&1&-2&-1\\0&0&0&h-5\end{bmatrix}
$$

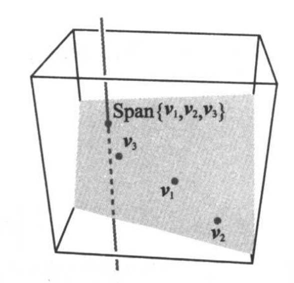


图 1-18　$h=5$ 时，点$\begin{bmatrix}-4\\3\\h\end{bmatrix}$位于与平面相交的直线上

该方程组相容当且仅当第 4 列没有主元. 即，$h-5$ 必须为0. 因此 $\boldsymbol{y}$ 在 Span$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$中当且仅当 $h=5$.（见图 1-18）

提醒：方程组中自由变量的出现并不能保证方程组是相容的.

1.4　矩阵方程 $Ax=b$

把向量的线性组合看作是向量与矩阵的乘积，这是线性代数的一个基本思想.

40

下述定义使我们可以把 1.3 节中的部分概念重述为新的形式.

【定义】　如果 A 为一个 $m\times n$ 矩阵，列为 $a_1,\cdots,a_n$，如果 $x\in\mathbf{R}^n$，则 A 与 x 的乘积是以 x 中对应元素为权重的 A 中列的线性组合，记作 Ax. 即，

$$Ax=[a_1\quad a_2\quad\cdots\quad a_n]\begin{bmatrix}x_1\\\vdots\\x_n\end{bmatrix}=x_1a_1+x_2a_2+\cdots+x_na_n$$

注意，仅当 A 的列数等于 x 中元素的个数时，Ax 才有意义.

【例题 1】

a. $\begin{bmatrix}1&2&-1\\0&-5&3\end{bmatrix}\begin{bmatrix}4\\3\\7\end{bmatrix}=4\begin{bmatrix}1\\0\end{bmatrix}+3\begin{bmatrix}2\\-5\end{bmatrix}+7\begin{bmatrix}-1\\3\end{bmatrix}=\begin{bmatrix}4\\0\end{bmatrix}+\begin{bmatrix}6\\-15\end{bmatrix}+\begin{bmatrix}-7\\21\end{bmatrix}=\begin{bmatrix}3\\6\end{bmatrix}$

b. $\begin{bmatrix}2&-3\\8&0\\-5&2\end{bmatrix}\begin{bmatrix}4\\7\end{bmatrix}=4\begin{bmatrix}2\\8\\-5\end{bmatrix}+7\begin{bmatrix}-3\\0\\2\end{bmatrix}=\begin{bmatrix}8\\32\\-20\end{bmatrix}+\begin{bmatrix}-21\\0\\14\end{bmatrix}=\begin{bmatrix}-13\\32\\-6\end{bmatrix}$ ■

【例题 2】　对 v_1，v_2，$v_3\in\mathbf{R}^m$，将线性组合 $3v_1-5v_2+7v_3$ 写成向量乘矩阵的形式.

解： 将 v_1，v_2，v_3 作为矩阵 A 的列，将权重 3，-5 和 7 作为向量 x 的元素. 即

$$3v_1-5v_2+7v_3=[v_1\quad v_2\quad v_3]\begin{bmatrix}3\\-5\\7\end{bmatrix}=Ax$$ ■

在 1.3 节中，我们学习了怎样把一个线性方程组写成向量方程形式. 例如，方程组

$$\begin{aligned}x_1+2x_2-x_3&=4\\-5x_2+3x_3&=1\end{aligned}\qquad(1)$$

等价于

41

$$x_1\begin{bmatrix}1\\0\end{bmatrix}+x_2\begin{bmatrix}2\\-5\end{bmatrix}+x_3\begin{bmatrix}-1\\3\end{bmatrix}=\begin{bmatrix}4\\1\end{bmatrix}\qquad(2)$$

如例题 2，我们可以把左端的线性组合写成向量乘矩阵的形式，所以(2)变成

$$\begin{bmatrix}1&2&-1\\0&-5&3\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}4\\1\end{bmatrix}\qquad(3)$$

方程(3)有 $Ax=b$ 的形式，为了区别于(2)中所示的向量方程，我们称这种形式的方程为**矩阵方程**(matrix equation).

注意，(3)中的矩阵就是方程组(1)的系数矩阵. 类似的计算表明，任意的线性方

程组，以及任意形如(2)的向量方程，都可以写成与之等价的矩阵方程形式 $\boldsymbol{Ax}=\boldsymbol{b}$. 这个简单的结论将在全书中反复使用.

正式的结论陈述如下：

【定理3】 设 A 是一个 $m\times n$ 矩阵，其中列为 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n,\boldsymbol{b}\in\mathbf{R}^m$，矩阵方程

$$\boldsymbol{Ax}=\boldsymbol{b} \tag{4}$$

与下列向量方程有相同的解集

$$x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\cdots+x_n\boldsymbol{a}_n=\boldsymbol{b} \tag{5}$$

从而与下列增广矩阵对应的线性方程组有相同的解集

$$[\boldsymbol{a}_1\quad \boldsymbol{a}_2\quad \cdots\quad \boldsymbol{a}_n\quad \boldsymbol{b}] \tag{6}$$

定理3为我们理解线性代数问题提供了一个有力的工具，因为现在可以把一个线性方程组看成三种不同但却相互等价的形式：矩阵方程、向量方程和线性方程组. 在现实生活中构造一个数学模型时，我们可以在任何情况下自由选择其中任何一种最自然、最便利的陈述形式. 矩阵方程、向量方程和线性方程组都用同一种方法求解——对增广矩阵(6)进行行化简. 另一种求解方法将在以后讨论.

1.4.1　解的存在性

由 $\boldsymbol{Ax}$ 的定义可直接得出下面这个有用的结论.

方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解当且仅当 $\boldsymbol{b}$ 为 A 中列的线性组合.

42

在1.3节中，我们考虑了存在性问题，"$\boldsymbol{b}$ 是否属于 $\mathrm{Span}\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3\}$?" 等价于 "$\boldsymbol{Ax}=\boldsymbol{b}$ 是否相容?" 接下来，我们将解决一个更难的存在性问题，即，方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 是否对 $\boldsymbol{b}$ 的所有取值都相容.

【例题3】 设 $A=\begin{bmatrix}1 & 3 & 4\\ -4 & 2 & -6\\ -3 & -2 & -7\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}b_1\\ b_2\\ b_3\end{bmatrix}$. 方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 是否对 b_1，b_2 和 b_3 的所有取值都相容?

解：对 $\boldsymbol{Ax}=\boldsymbol{b}$ 的增广矩阵进行行化简：

$$\begin{bmatrix}1 & 3 & 4 & b_1\\ -4 & 2 & -6 & b_2\\ -3 & -2 & -7 & b_3\end{bmatrix}\sim\begin{bmatrix}1 & 3 & 4 & b_1\\ 0 & 14 & 10 & b_2+4b_1\\ 0 & 7 & 5 & b_3+3b_1\end{bmatrix}$$

$$\sim\begin{bmatrix}1 & 3 & 4 & b_1\\ 0 & 14 & 10 & b_2+4b_1\\ 0 & 0 & 0 & b_3+3b_1-\frac{1}{2}(b_2+4b_1)\end{bmatrix}$$

增广列中第三个元素为 $b_1-\frac{1}{2}b_2+b_3$. 因为 $\boldsymbol{b}$ 取某些值时 $b_1-\frac{1}{2}b_2+b_3$ 不等于0，所以方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 并不是对 $\boldsymbol{b}$ 的所有取值都相容. ■

例题3中的简化矩阵表示出了所有使 $\boldsymbol{Ax}=\boldsymbol{b}$ 相容的 $\boldsymbol{b}$：$\boldsymbol{b}$ 中的元素必须满足

$$b_1 - \frac{1}{2}b_2 + b_3 = 0$$

这是 $\mathbf{R}^3$ 中过原点的一个平面的方程．该平面为 $\boldsymbol{A}$ 中三列的所有线性组合构成的集合．见图 1-19．

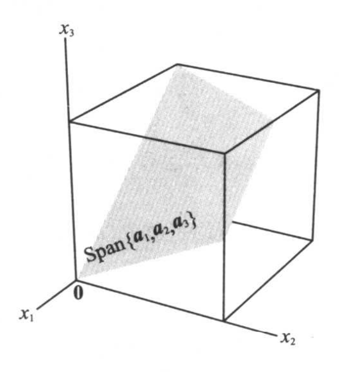


图 1-19　$\boldsymbol{A} = [\boldsymbol{a}_1\ \boldsymbol{a}_2\ \boldsymbol{a}_3]$ 中的列生成过 $\mathbf{0}$ 点的平面

因为 $\boldsymbol{A}$ 的阶梯形式有全为 0 的行，所以例题 3 中的方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 不能对 $\boldsymbol{b}$ 的所有取值都相容．如果 $\boldsymbol{A}$ 中的三行都有主元，则增广矩阵的阶梯形式不可能有形如[0　0　0　1]的行，在这种情况下就不用担心增广列的计算．

在下面的定理中，我们说 **$\boldsymbol{A}$ 的列张成 $\mathbf{R}^m$**，是指任意 $\boldsymbol{b} \in \mathbf{R}^m$ 都是 $\boldsymbol{A}$ 中列的线性组合．通常说 $\mathbf{R}^m$ 中的一个向量集 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$**张成**(span)或**生成**(generate) $\mathbf{R}^m$，是指 $\mathbf{R}^m$ 中的每个向量都是 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 的一个线性组合，即 $\mathrm{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\} = \mathbf{R}^m$．

【定理 4】　设 $\boldsymbol{A}$ 是 $m \times n$ 矩阵，则下列命题等价．即，对于指定的 $\boldsymbol{A}$，下列命题或者同为真或者同为假．

a. 对任意 $\boldsymbol{b} \in \mathbf{R}^m$，方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 都有解．
b. 任意 $\boldsymbol{b} \in \mathbf{R}^m$ 都是 $\boldsymbol{A}$ 中列的一个线性组合．
c. $\boldsymbol{A}$ 的列张成 $\mathbf{R}^m$．
d. $\boldsymbol{A}$ 中每一行都有主元位置．

43

定理 4 是本章最有用的定理之一．根据 $\boldsymbol{Ax}$ 的定义以及向量集张成 $\mathbf{R}^m$ 的含义，(a)、(b)和(c)等价．例题 3 后面的讨论说明了(a)和(b)等价的原因；本节末尾将给出一个证明．习题中给出了应用定理 4 的若干例子．

警告：定理 4 与系数矩阵而非增广矩阵有关．如果增广矩阵$[\boldsymbol{A}\ \boldsymbol{b}]$的每一行中都含有主元位置，方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 的相容性仍无法确定．

1.4.2　*Ax* 的计算

例题 1 的计算应用了矩阵 $\boldsymbol{A}$ 与向量 $\boldsymbol{x}$ 乘积的定义．而下面这个简单例题将介绍一种更有效的计算 $\boldsymbol{Ax}$ 的方法．

【例题 4】　计算 $\boldsymbol{Ax}$，$\boldsymbol{A} = \begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}$，$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$．

解： 由定义得，

$$\begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1\begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 5 \\ -2 \end{bmatrix} + x_3\begin{bmatrix} 4 \\ -3 \\ 8 \end{bmatrix}$$

$$= \begin{bmatrix} 2x_1 \\ -x_1 \\ 6x_1 \end{bmatrix} + \begin{bmatrix} 3x_2 \\ 5x_2 \\ -2x_2 \end{bmatrix} + \begin{bmatrix} 4x_3 \\ -3x_3 \\ 8x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ -x_1 + 5x_2 - 3x_3 \\ 6x_1 - 2x_2 + 8x_3 \end{bmatrix} \tag{7}$$

乘积 $\boldsymbol{Ax}$ 中的第一个元素是若干乘积(有时也称作点乘)之和，它们分别是 $\boldsymbol{A}$ 中第 1 行元素与 $\boldsymbol{x}$ 中对应元素的积，即

$$\begin{bmatrix} 2 & 3 & 4 \\ & & \\ & & \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 2x_1+3x_2+4x_3 \\ \\ \end{bmatrix}$$

这里没写出(7)的全部计算步骤，但说明了怎样直接计算 $\boldsymbol{Ax}$ 的第一个元素. 类似地，$\boldsymbol{Ax}$ 的第二个元素也可以用 $\boldsymbol{A}$ 中第 2 行元素乘以 $\boldsymbol{x}$ 中对应元素，再对乘积求和而得到：

$$\begin{bmatrix} & & \\ -1 & 5 & -3 \\ & & \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} \\ -x_1+5x_2-3x_3 \\ \\ \end{bmatrix}$$

44

同样地，$\boldsymbol{Ax}$ 的第三个元素也可以由 $\boldsymbol{A}$ 的第 3 行与 $\boldsymbol{x}$ 的元素计算得出. ■

计算 $\boldsymbol{Ax}$ 的行 - 向量法则

若乘积 $\boldsymbol{Ax}$ 有定义，则 $\boldsymbol{Ax}$ 的第个 i 元素是 $\boldsymbol{A}$ 中第 i 行元素与 $\boldsymbol{x}$ 中对应元素的乘积之和.

【例题 5】

(a) $\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix}=\begin{bmatrix} 1\times4+2\times3+(-1)\times7 \\ 0\times4+(-5)\times3+3\times7 \end{bmatrix}=\begin{bmatrix} 3 \\ 6 \end{bmatrix}$

(b) $\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix}\begin{bmatrix} 4 \\ 7 \end{bmatrix}=\begin{bmatrix} 2\times4+(-3)\times7 \\ 8\times4+0\times7 \\ (-5)\times4+2\times7 \end{bmatrix}=\begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} r \\ s \\ t \end{bmatrix}=\begin{bmatrix} 1\times r+0\times s+0\times t \\ 0\times r+1\times s+0\times t \\ 0\times r+0\times s+1\times t \end{bmatrix}=\begin{bmatrix} r \\ s \\ t \end{bmatrix}$ ■

根据定义，例题 5(c) 中的矩阵对角线上都是 1，其他位置均为 0，该矩阵称为**单位矩阵**(identity matrix)，记为 $\boldsymbol{I}$. (c) 中的计算表明，对任意 $\boldsymbol{x}\in\mathbf{R}^3$ 都有 $\boldsymbol{Ix}=\boldsymbol{x}$. 类似地，存在 $n\times n$ 单位矩阵，记成 $\boldsymbol{I}_n$. 与(c)中一样，对任意 $\boldsymbol{x}\in\mathbf{R}^n$ 都有 $\boldsymbol{I}_n\boldsymbol{x}=\boldsymbol{x}$.

1.4.3 矩阵 - 向量乘积 Ax 的性质

下面的定理很重要，在全书中都用到. 定理的证明依赖于 $\boldsymbol{Ax}$ 的定义和 $\mathbf{R}^n$ 的代数性质.

【定理 5】 如果 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，$\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是 $\mathbf{R}^n$ 中的向量，c 是一个数量，则：

a. $\boldsymbol{A}(\boldsymbol{u}+\boldsymbol{v})=\boldsymbol{Au}+\boldsymbol{Av}$;

b. $\boldsymbol{A}(c\boldsymbol{u})=c(\boldsymbol{Au})$.

证明： 为了简化证明，设 $n=3$，$\boldsymbol{A}=[\boldsymbol{a}_1\quad\boldsymbol{a}_2\quad\boldsymbol{a}_3]$，$\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是 $\mathbf{R}^3$ 中的向量. (一般情况的证明也类似.) 设 u_i 和 v_i 分别为 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 中的第 i 个元素，$i=1,2,3$. 为证明(a)，

45 计算 $A(u+v)$，将它看成以 $u+v$ 中元素为权重的 A 中列的线性组合.

$$A(u+v)=[a_1\quad a_2\quad a_3]\begin{bmatrix}u_1+v_1\\u_2+v_2\\u_3+v_3\end{bmatrix}$$

$u+v$中的元素

$$=(u_1+v_1)a_1+(u_2+v_2)a_2+(u_3+v_3)a_3$$

A中的列

$$=(u_1a_1+u_2a_2+u_3a_3)+(v_1a_1+v_2a_2+v_3a_3)$$
$$=Au+Av$$

为证明(b)，计算 $A(cu)$，将它看成以 cu 中元素为权重的 A 中列的线性组合.

$$A(cu)=[a_1\quad a_2\quad a_3]\begin{bmatrix}cu_1\\cu_2\\cu_3\end{bmatrix}$$
$$=(cu_1)a_1+(cu_2)a_2+(cu_3)a_3$$
$$=c(u_1a_1)+c(u_2a_2)+c(u_3a_3)$$
$$=c(u_1a_1+u_2a_2+u_3a_3)$$
$$=c(Au)$$ ■

注记

为了优化 Ax 的计算机算法，需要把数据存储在连续的内存单元里. 应用得最广泛的专业矩阵计算程序是用 Fortran 语言编写的，它把矩阵存储成列的集合，因此计算 Ax 即是计算 A 中列的线性组合. 而如果程序使用较流行的 C 语言来编写，则矩阵按行存储，那么计算 Ax 就需要使用行-向量法则.

定理4的证明：从定理4后面的讨论知道(a)、(b)和(c)等价. 因此只要证明(对任意的矩阵 A)(a)和(d)同为真或者同为假，四个命题就互相等价了.

设 U 为 A 的阶梯形式. 给定 $b\in\mathbf{R}^m$，我们对增广矩阵 $[A\ b]$ 进行行化简，得到 $[U\ d]$，$d\in\mathbf{R}^m$：

$$[A\ b]\sim\cdots\sim[U\ d]$$

如果(d)是正确的，则 U 的每一行都包含主元位置，且增广列中不含主元位置. 因此 $Ax=b$ 对任意的 b 都有解，于是(a)是正确的. 如果(d)不正确，则 U 的最后一行全部为0. 设 d 为最后一个元素为1的任意向量. 则 $[U\ d]$ 表示的是一个不相容的方程组. 由于行变换可逆，$[U\ d]$ 可以变回 $[A\ b]$，得到的方程组 $Ax=b$ 仍然不相容，因此(a)不正确. ■

46

基础练习

1. 设 $A=\begin{bmatrix}1&5&-2&0\\-3&1&9&-5\\4&-8&-1&7\end{bmatrix}$，$p=\begin{bmatrix}3\\-2\\0\\-4\end{bmatrix}$，$b=\begin{bmatrix}-7\\9\\0\end{bmatrix}$. 证明 p 是 $Ax=b$ 的一个解. 利用这个结论将 b 表示为 A 中列的线性组合.

2. 设 $\boldsymbol{A}=\begin{bmatrix}2&5\\3&1\end{bmatrix}$，$\boldsymbol{u}=\begin{bmatrix}4\\-1\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}-3\\5\end{bmatrix}$. 通过计算 $\boldsymbol{A}(\boldsymbol{u}+\boldsymbol{v})$ 和 $\boldsymbol{Au}+\boldsymbol{Av}$ 来验证定理 5(a).

习题 1.4

利用例题 1 中(a)以及(b)的计算 $\boldsymbol{Ax}$ 的行－向量法则，计算习题 1 ~ 4 中的乘积. 若某个乘积无定义，试解释原因.

1. $\begin{bmatrix}-4&2\\1&6\\0&1\end{bmatrix}\begin{bmatrix}3\\-2\\7\end{bmatrix}$　　2. $\begin{bmatrix}2\\6\\-1\end{bmatrix}\begin{bmatrix}5\\-1\end{bmatrix}$　　3. $\begin{bmatrix}6&5\\-4&-3\\7&6\end{bmatrix}\begin{bmatrix}2\\-3\end{bmatrix}$　　4. $\begin{bmatrix}8&3&-4\\5&1&2\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}$

在习题 5 ~ 8 中，利用 $\boldsymbol{Ax}$ 的定义把矩阵方程写成向量方程的形式，或把向量方程写成矩阵方程的形式.

5. $\begin{bmatrix}5&1&-8&4\\-2&-7&3&-5\end{bmatrix}\begin{bmatrix}5\\-1\\3\\-2\end{bmatrix}=\begin{bmatrix}-8\\16\end{bmatrix}$　　6. $\begin{bmatrix}7&-3\\2&1\\9&-6\\-3&2\end{bmatrix}\begin{bmatrix}-2\\-5\end{bmatrix}=\begin{bmatrix}1\\-9\\12\\-4\end{bmatrix}$

7. $x_1\begin{bmatrix}4\\-1\\7\\-4\end{bmatrix}+x_2\begin{bmatrix}-5\\3\\-5\\1\end{bmatrix}+x_3\begin{bmatrix}7\\-8\\0\\2\end{bmatrix}=\begin{bmatrix}6\\-8\\0\\-7\end{bmatrix}$　　8. $z_1\begin{bmatrix}4\\-2\end{bmatrix}+z_2\begin{bmatrix}-4\\5\end{bmatrix}+z_3\begin{bmatrix}-5\\4\end{bmatrix}+z_4\begin{bmatrix}3\\0\end{bmatrix}=\begin{bmatrix}4\\13\end{bmatrix}$

在习题 9 和 10 中，先把方程组写成向量方程，然后写成矩阵方程.

9. $\begin{aligned}3x_1+x_2-5x_3&=9\\x_2+4x_3&=0\end{aligned}$　　10. $\begin{aligned}8x_1-x_2&=4\\5x_1+4x_2&=1\\x_1-3x_2&=2\end{aligned}$

习题 11 和 12 中给出了 $\boldsymbol{A}$ 和 $\boldsymbol{b}$，写出与矩阵方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 对应的线性方程组的增广矩阵. 求解方程组，并把解写成向量形式.

11. $\boldsymbol{A}=\begin{bmatrix}1&2&4\\0&1&5\\-2&-4&-3\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}-2\\2\\9\end{bmatrix}$　　12. $\boldsymbol{A}=\begin{bmatrix}1&2&1\\-3&-1&2\\0&5&3\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}0\\1\\-1\end{bmatrix}$

13. 设 $\boldsymbol{u}=\begin{bmatrix}0\\4\\4\end{bmatrix}$，$\boldsymbol{A}=\begin{bmatrix}3&-5\\-2&6\\1&1\end{bmatrix}$，$\boldsymbol{u}$ 是否在 $\mathbf{R}^3$ 中由 $\boldsymbol{A}$ 的列张成的平面里(见右图)？为什么？

14. 设 $\boldsymbol{u}=\begin{bmatrix}2\\-3\\2\end{bmatrix}$，$\boldsymbol{A}=\begin{bmatrix}5&8&7\\0&1&-1\\1&3&0\end{bmatrix}$. $\boldsymbol{u}$ 是否在 $\mathbf{R}^3$ 中由 $\boldsymbol{A}$ 的列张成的子集里？为什么？

u?
u?
u在哪里?

习题 13 图

15. 设 $\boldsymbol{A}=\begin{bmatrix}2&-1\\-6&3\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}b_1\\b_2\end{bmatrix}$. 请说明方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 并不是对任意的 $\boldsymbol{b}$ 都有解，并且表示出使 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解的所有 $\boldsymbol{b}$ 构成的集合.

16. 对 $A=\begin{bmatrix}1&-3&-4\\-3&2&6\\5&-1&-8\end{bmatrix}$, $b=\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$重做习题15.

习题17～20中的矩阵 A 和 B 如下. 利用适当的定理和计算来给出解答.

$$A=\begin{bmatrix}1&3&0&3\\-1&-1&-1&1\\0&-4&2&-8\\2&0&3&-1\end{bmatrix}\quad B=\begin{bmatrix}1&3&-2&2\\0&1&1&-5\\1&2&-3&7\\-2&-8&2&-1\end{bmatrix}$$

17. A 中有多少行包含主元位置? 是否对 $\mathbf{R}^4$ 中的任意 b, $Ax=b$ 都有解?
18. B 的列是否张成 $\mathbf{R}^4$? 是否对 $\mathbf{R}^4$ 中的任意 y, $Bx=y$ 都有解?
19. $\mathbf{R}^4$ 中的每个向量是否都可以写成 A 中列的线性组合形式? A 的列是否张成 $\mathbf{R}^4$?
20. $\mathbf{R}^4$ 中的每个向量是否都可以写成 B 中列的线性组合形式? B 的列是否张成 $\mathbf{R}^3$?
21. 设 $v_1=\begin{bmatrix}1\\0\\-1\\0\end{bmatrix}$, $v_2=\begin{bmatrix}0\\-1\\0\\1\end{bmatrix}$, $v_3=\begin{bmatrix}1\\0\\0\\-1\end{bmatrix}$. $\{v_1,v_2,v_3\}$是否张成 $\mathbf{R}^4$? 为什么?
22. 设 $v_1=\begin{bmatrix}0\\0\\-2\end{bmatrix}$, $v_2=\begin{bmatrix}0\\-3\\8\end{bmatrix}$, $v_3=\begin{bmatrix}4\\-1\\-5\end{bmatrix}$. $\{v_1,v_2,v_3\}$是否张成 $\mathbf{R}^3$? 为什么?

在习题23和24中, 判断各命题的真假, 并说明理由.

23. a. $Ax=b$ 是向量方程.
    b. 向量 b 是矩阵 A 中列的线性组合当且仅当方程 $Ax=b$ 至少有一个解.
    c. 如果增广矩阵$[A\ b]$的每一行都有一个主元位置, 则方程 $Ax=b$ 相容.
    d. 乘积 Ax 中的第一个元素是一些乘积的和.
    e. 如果 $m\times n$ 矩阵 A 的列张成 $\mathbf{R}^m$, 则方程 $Ax=b$ 对于 $\mathbf{R}^m$ 中任意 b 相容.
    f. 如果 A 是一个 $m\times n$ 矩阵, 且对于 $\mathbf{R}^m$ 中某些 b, $Ax=b$ 相容, 则不是 A 的每一行都含有主元位置.
24. a. 每个矩阵方程 $Ax=b$ 都对应于一个同解的向量方程.
    b. 如果给出适当的 A 和 x, 所有的向量线性组合都可以写成 Ax 的形式.
    c. 如果 $A=[a_1\ \ a_2\ \ a_3]$, 则增广矩阵为$[a_1\ \ a_2\ \ a_3\ \ b]$的线性方程组与 $Ax=b$ 的解集相同.
    d. 如果方程 $Ax=b$ 相容, 则 b 不在 A 的列张成的集合中.
    e. 如果增广矩阵$[A\ b]$的每一行都含有一个主元位置, 则方程 $Ax=b$ 不相容.
    f. 如果 A 是一个 $m\times n$ 矩阵, 并且它的列没有张成 $\mathbf{R}^m$, 则存在 $\mathbf{R}^m$ 中的 b, 使方程 $Ax=b$ 不相容.
25. 易知$\begin{bmatrix}4&-3&1\\5&-2&5\\-6&2&-3\end{bmatrix}\begin{bmatrix}-3\\-1\\2\end{bmatrix}=\begin{bmatrix}-7\\-3\\10\end{bmatrix}$. 利用这个等式(不进行行变换)求出数量 c_1, c_2, c_3, 使得

$$\begin{bmatrix}-7\\-3\\10\end{bmatrix}=c_1\begin{bmatrix}4\\5\\-6\end{bmatrix}+c_2\begin{bmatrix}-3\\-2\\2\end{bmatrix}+c_3\begin{bmatrix}1\\5\\-3\end{bmatrix}.$$

26. 设 $\boldsymbol{u}=\begin{bmatrix}7\\2\\5\end{bmatrix}$, $\boldsymbol{v}=\begin{bmatrix}3\\1\\3\end{bmatrix}$, $\boldsymbol{w}=\begin{bmatrix}6\\1\\0\end{bmatrix}$. 易知 $3\boldsymbol{u}-5\boldsymbol{v}-\boldsymbol{w}=\boldsymbol{0}$. 用这个结果(不进行行变换)求出数量$x_1$, x_2, 满足方程

$$\begin{bmatrix}7&3\\2&1\\5&3\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}6\\1\\0\end{bmatrix}.$$

27. 设 $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, $\boldsymbol{q}_3$和 $\boldsymbol{v}$ 表示 $\mathbf{R}^5$中的向量, x_1, x_2, x_3表示数量. 把下面的向量方程改写成矩阵方程. 使用的符号要保持前后一致.

$$x_1\boldsymbol{q}_1+x_2\boldsymbol{q}_2+x_3\boldsymbol{q}_3=\boldsymbol{v}$$

28. 用向量方程中的符号形式改写下面的矩阵方程, 用 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots$表示向量, $c_1,c_2,\cdots$表示数量. 并用矩阵方程中的数据解释向量方程中的符号.

48

$$\begin{bmatrix}-3&5&-4&9&7\\5&8&1&-2&-4\end{bmatrix}\begin{bmatrix}-3\\2\\4\\-1\\2\end{bmatrix}=\begin{bmatrix}8\\-1\end{bmatrix}$$

29. 构造一个非阶梯形式的 3×3 矩阵, 使它的列张成 $\mathbf{R}^3$. 并说明构造的矩阵具备要求的性质.
30. 构造一个非阶梯形式的3×3矩阵, 使它的列不能张成 $\mathbf{R}^3$. 并说明构造的矩阵具备要求的性质.
31. 设 A 是一个 3×2 矩阵. 为什么方程 $A\boldsymbol{x}=\boldsymbol{b}$ 不能对 $\mathbf{R}^3$中的任意 $\boldsymbol{b}$ 都相容? 将此结论推广到对任意行数多于列数的矩阵 A.
32. $\mathbf{R}^4$中三个向量构成的集合是否能张成 $\mathbf{R}^4$? 说明理由. $\mathbf{R}^m$中 $n(n<m)$个向量构成的集合是否能张成 $\mathbf{R}^m$?
33. 设 A 是一个 4×3 矩阵, $\boldsymbol{b}$ 为 $\mathbf{R}^4$中的向量, 且 $A\boldsymbol{x}=\boldsymbol{b}$ 有唯一解. 你能得出什么关于 A 的简化阶梯形式的结论? 给出证明.
34. 设 A 是一个 3×3 矩阵, $\boldsymbol{b}$ 为 $\mathbf{R}^3$中的向量, 且 $A\boldsymbol{x}=\boldsymbol{b}$ 有唯一解. 为什么 A 中的列一定能张成 $\mathbf{R}^3$? 给出理由.
35. 设 A 是一个 3×4 矩阵, $\boldsymbol{y}_1$和 $\boldsymbol{y}_2$是 $\mathbf{R}^3$中的向量, 设 $\boldsymbol{w}=\boldsymbol{y}_1+\boldsymbol{y}_2$. 假设存在 $\mathbf{R}^4$中的 $\boldsymbol{x}_1$和 $\boldsymbol{x}_2$使得 $\boldsymbol{y}_1=A\boldsymbol{x}_1$, $\boldsymbol{y}_2=A\boldsymbol{x}_2$. 如何证明方程组 $A\boldsymbol{x}=\boldsymbol{w}$ 相容? (注意: $\boldsymbol{x}_1$和 $\boldsymbol{x}_2$表示向量, 而不是向量中的元素.)
36. 设 A 是一个 5×3 矩阵, $\boldsymbol{y}$ 是 $\mathbf{R}^3$中向量, $\boldsymbol{z}$ 是 $\mathbf{R}^5$中向量. 假设 $A\boldsymbol{y}=\boldsymbol{z}$. 什么条件能使 $A\boldsymbol{x}=4\boldsymbol{z}$ 相容?

[M]在习题 37 ~ 40 中, 确定矩阵的列能否张成 $\mathbf{R}^4$.

37. $\begin{bmatrix}7&2&-5&8\\-5&-3&4&-9\\6&10&-2&7\\-7&9&2&15\end{bmatrix}$

38. $\begin{bmatrix}5&-7&-4&9\\6&-8&-7&5\\4&-4&-9&-9\\-9&11&16&7\end{bmatrix}$

39. $\begin{bmatrix}12&-7&11&-9&5\\-9&4&-8&7&-3\\-6&11&-7&3&-9\\4&-6&10&-5&12\end{bmatrix}$

40. $\begin{bmatrix}8&11&-6&-7&13\\-7&-8&5&6&-9\\11&7&-7&-9&-6\\-3&4&1&8&7\end{bmatrix}$

41. [M]找出习题39中矩阵可以删除的一列，使余下的列仍然能张成 $\mathbf{R}^4$.

42. [M]找出习题40中矩阵可以删除的一列，使余下的列仍然能张成 $\mathbf{R}^4$. 可以删除更多的列吗？

SG 掌握线性代数的概念：张成 1 ~ 19　　　CD 解 $A\boldsymbol{x}=\boldsymbol{b}$.

基础练习答案

1. 矩阵方程

$$\begin{bmatrix} 1 & 5 & -2 & 0 \\ -3 & 1 & 9 & -5 \\ 4 & -8 & -1 & 7 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \\ 0 \\ -4 \end{bmatrix}=\begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}$$

等价于向量方程

$$3\begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix}-2\begin{bmatrix} 5 \\ 1 \\ -8 \end{bmatrix}+0\begin{bmatrix} -2 \\ 9 \\ -1 \end{bmatrix}-4\begin{bmatrix} 0 \\ -5 \\ 7 \end{bmatrix}=\begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}$$

49

该向量方程把 $\boldsymbol{b}$ 表示成 $\boldsymbol{A}$ 中列的线性组合形式.

2. $\boldsymbol{u}+\boldsymbol{v}=\begin{bmatrix} 4 \\ -1 \end{bmatrix}+\begin{bmatrix} -3 \\ 5 \end{bmatrix}=\begin{bmatrix} 1 \\ 4 \end{bmatrix}$

$$A(\boldsymbol{u}+\boldsymbol{v})=\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \end{bmatrix}=\begin{bmatrix} 2+20 \\ 3+4 \end{bmatrix}=\begin{bmatrix} 22 \\ 7 \end{bmatrix}$$

$$A\boldsymbol{u}+A\boldsymbol{v}=\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ -1 \end{bmatrix}+\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} -3 \\ 5 \end{bmatrix}=\begin{bmatrix} 3 \\ 11 \end{bmatrix}+\begin{bmatrix} 19 \\ -4 \end{bmatrix}=\begin{bmatrix} 22 \\ 7 \end{bmatrix}$$

1.5　线性方程组的解集

线性方程组的解集是线性代数研究的一个重要对象，后面章节中都将涉及. 本节利用向量的记号给出解集的显式几何表示.

1.5.1　齐次线性方程组

如果一个线性方程组可以写成 $A\boldsymbol{x}=\boldsymbol{0}$ 的形式，则称该方程组是**齐次**(homogeneous)的，这里的 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，$\boldsymbol{0}$ 是 $\mathbf{R}^m$ 中的零向量. 这样的方程组至少含有一个解，即 $\boldsymbol{x}=\boldsymbol{0}$($\mathbf{R}^n$ 中的零向量). 该零解通常称为**平凡解**(trivial solution). 对于给定的方程 $A\boldsymbol{x}=\boldsymbol{0}$，最重要的问题是是否存在**非平凡解**(nontrivial solution)，即，是否存在非零向量 $\boldsymbol{x}$ 满足 $A\boldsymbol{x}=\boldsymbol{0}$. 由1.2节的存在性和唯一性定理可直接得到如下结论.

齐次方程 $A\boldsymbol{x}=\boldsymbol{0}$ 有非平凡解当且仅当方程至少含有一个自由变量.

【例题1】　判断下列齐次方程是否有非平凡解. 表示其解集.

$$\begin{aligned} 3x_1+5x_2-4x_3&=0 \\ -3x_1-2x_2+4x_3&=0 \\ 6x_1+x_2-8x_3&=0 \end{aligned}$$

解： 设 $\boldsymbol{A}$ 为方程组的系数矩阵. 对增广矩阵 $[\boldsymbol{A}\ \boldsymbol{0}]$ 进行行化简得到阶梯形式：

$$\begin{bmatrix} 3 & 5 & -4 & 0 \\ -3 & -2 & 4 & 0 \\ 6 & 1 & -8 & 0 \end{bmatrix} \sim \begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & -9 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

50

因为 x_3是自由变量，则 $A\boldsymbol{x}=\boldsymbol{0}$ 有非平凡解(对应于 x_3的不同取值). 要表示出解集，则继续对$[A\ \boldsymbol{0}]$进行行化简得到简化阶梯形式：

$$\begin{bmatrix} 1 & 0 & -\frac{4}{3} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{aligned} x_1 \quad - \tfrac{4}{3}x_3 &= 0 \\ x_2 \qquad &= 0 \\ 0 &= 0 \end{aligned}$$

求出基本变量 $x_1=\frac{4}{3}x_3$，$x_2=0$，其中 x_3是自由变量. $A\boldsymbol{x}=\boldsymbol{0}$ 的通解写成向量有如下形式：

$$\boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix}=x_3\begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}=x_3\boldsymbol{v}, \text{其中 } \boldsymbol{v}=\begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}$$

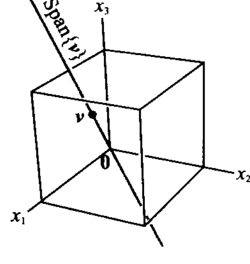


图 1-20

这里 x_3作为因子被提取出来. 通解的这个向量表示说明，$A\boldsymbol{x}=\boldsymbol{0}$ 的每个解都是 $\boldsymbol{v}$ 的数量倍数. 取 $x_3=0$ 时即得到平凡解. 从几何上看，该解集是 $\mathbf{R}^3$ 中过 $\boldsymbol{0}$ 点的一条直线. 见图 1-20.

■

注意，非平凡解 $\boldsymbol{x}$ 中也可以有零元素，只要元素不全为零即可.

【例题 2】 单个线性方程可以看作一个简单的线性方程组. 表示出下面齐次“方程组”的全体解.

$$10x_1-3x_2-2x_3=0 \tag{1}$$

解：无需使用矩阵记号，即可用自由变量表示出基本变量 x_1，得到通解 $x_1=0.3x_2+0.2x_3$，其中 x_2和 x_3为自由变量. 写成向量为

$$\begin{aligned} \boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}&=\begin{bmatrix} 0.3x_2+0.2x_3 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0.3x_2 \\ x_2 \\ 0 \end{bmatrix}+\begin{bmatrix} 0.2x_3 \\ 0 \\ x_3 \end{bmatrix} \\ &=x_2\underset{\uparrow\, \boldsymbol{u}}{\begin{bmatrix} 0.3 \\ 1 \\ 0 \end{bmatrix}}+x_3\underset{\uparrow\, \boldsymbol{v}}{\begin{bmatrix} 0.2 \\ 0 \\ 1 \end{bmatrix}} \quad (x_2 \text{ 和 } x_3 \text{ 是自由变量}) \end{aligned} \tag{2}$$

上述计算过程表明，(1)的每个解都是向量 $\boldsymbol{u}$，$\boldsymbol{v}$ 的线性组合，如(2)中所示. 由于 $\boldsymbol{u}$

和 $\boldsymbol{v}$ 之间不存在数量倍数的关系，所以解集是过原点的一个平面，见图 1-21. ■

例题 1 和例题 2 以及部分课后习题，说明适当选取 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 时，齐次方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集总可以表示成 $\mathrm{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$. 51 如果零向量是唯一解，则解集为 $\mathrm{Span}\{\boldsymbol{0}\}$，如果方程 $A\boldsymbol{x}=\boldsymbol{0}$ 仅有一个自由变量，则解集为过原点的一条直线，如图 1-20 所示. 图 1-21 中过原点的平面则是当方程 $A\boldsymbol{x}=\boldsymbol{0}$ 含两个或两个以上自由变量时解集的示意图. 注意，即使 $\boldsymbol{u}$，$\boldsymbol{v}$ 不是 $A\boldsymbol{x}=\boldsymbol{0}$ 的解，我们也可以用这幅图表示 $\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$. 见 1.3 节图 1-17.

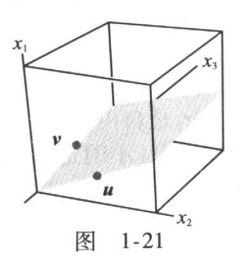


图　1-21

1.5.2　参数向量形式

例题 2 中平面的原始方程(1)给出了平面的隐式表示. 求解该方程就是求平面的显式表示，即把它写成 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的张成集的形式. 我们称方程(2)为平面的**参数向量方程**(parametric vector equation). 为了强调参数取遍所有的实数，这样的方程可以写为

$$\boldsymbol{x}=s\boldsymbol{u}+t\boldsymbol{v}(s,t\in\mathbf{R})$$

例题 1 中方程 $\boldsymbol{x}=x_3\boldsymbol{v}$($x_3$为自由变量)及 $\boldsymbol{x}=t\boldsymbol{v}(t\in\mathbf{R})$均为直线的参数向量方程. 只要将解集如例题 1，2 那样用向量明确地表示出来，我们就说解具有**参数向量形式**(parametric vector form).

1.5.3　非齐次方程组的解

当非齐次线性方程组有多个解时，通解可以写成参数向量形式，即，一个向量加上满足对应齐次方程组的若干向量的任意线性组合.

【例题 3】　求出方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的全体解

$$A=\begin{bmatrix}3&5&-4\\-3&-2&4\\6&1&-8\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}7\\-1\\-4\end{bmatrix}$$

解：A 是例题 1 中方程组的系数矩阵. 对$[A\ \boldsymbol{b}]$进行行变换得到

$$\begin{bmatrix}3&5&-4&7\\-3&-2&4&-1\\6&1&-8&-4\end{bmatrix}\sim\begin{bmatrix}1&0&-\frac{4}{3}&-1\\0&1&0&2\\0&0&0&0\end{bmatrix},\quad\begin{aligned}x_1\qquad -\tfrac{4}{3}x_3&=-1\\x_2\qquad\quad&=2\\0&=0\end{aligned}$$

因此，$x_1=-1+\frac{4}{3}x_3$，$x_2=2$，其中 x_3是自由变量. $A\boldsymbol{x}=\boldsymbol{b}$ 的通解写成向量有如下形式：

$$\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}-1+\frac{4}{3}x_3\\2\\x_3\end{bmatrix}=\begin{bmatrix}-1\\2\\0\end{bmatrix}+\begin{bmatrix}\frac{4}{3}x_3\\0\\x_3\end{bmatrix}=\underset{\uparrow\atop \boldsymbol{p}}{\begin{bmatrix}-1\\2\\0\end{bmatrix}}+x_3\underset{\uparrow\atop \boldsymbol{v}}{\begin{bmatrix}\frac{4}{3}\\0\\1\end{bmatrix}}$$

52

方程 $\boldsymbol{x}=\boldsymbol{p}+x_3\boldsymbol{v}$，或者用 t 表示一般参数，

$$\boldsymbol{x}=\boldsymbol{p}+t\boldsymbol{v}\quad(t\in\mathbf{R})\tag{3}$$

就是方程 $A\boldsymbol{x}=\boldsymbol{b}$ 解集的向量形式. 回忆例题 1 中 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集有参数向量方程形式

$$\boldsymbol{x}=t\boldsymbol{v}\quad(t\in\mathbf{R})\tag{4}$$

[其中的 $\boldsymbol{v}$ 与(3)中相同]. 因此 $A\boldsymbol{x}=\boldsymbol{b}$ 的解可以由向量 $\boldsymbol{p}$ 与 $A\boldsymbol{x}=\boldsymbol{0}$ 的解相加得到. 向量 $\boldsymbol{p}$ 本身是 $A\boldsymbol{x}=\boldsymbol{b}$ 的一个特解[对应于(3)中的 $t=0$]. ■

为了给出方程 $A\boldsymbol{x}=\boldsymbol{b}$ 解集的几何表示，我们可以考虑把向量加法看作平移. 给定 $\boldsymbol{v}$，$\boldsymbol{p}\in\mathbf{R}^2$ 或 $\mathbf{R}^3$，把 $\boldsymbol{p}$ 加到 $\boldsymbol{v}$ 上相当于沿着与过 $\boldsymbol{0}$ 和 $\boldsymbol{p}$ 的直线平行的方向移动 $\boldsymbol{v}$. 我们说 **$\boldsymbol{v}$ 平移 $\boldsymbol{p}$ 得到 $\boldsymbol{v}+\boldsymbol{p}$**. 见图 1-22. 如果将 $\mathbf{R}^2$ 或 $\mathbf{R}^3$ 中直线 L 上每一点都平移 $\boldsymbol{p}$，可以得到一条与 L 平行的直线. 见图 1-23.

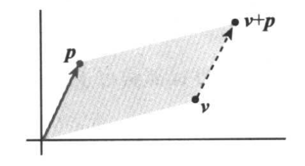


图 1-22 $\boldsymbol{v}$ 加上 $\boldsymbol{p}$ 得到 $\boldsymbol{v}+\boldsymbol{p}$

假设 L 是一条过 $\boldsymbol{0}$ 和 $\boldsymbol{v}$ 的直线，用方程(4)表示. 在 L 的每一点上加 $\boldsymbol{p}$，平移后所得的直线方程是(3). 注意 $\boldsymbol{p}$ 在直线(3)上. 我们称(3)为**过 $\boldsymbol{p}$ 点平行于 $\boldsymbol{v}$ 的直线方程**. 因此 $A\boldsymbol{x}=\boldsymbol{b}$ 的解集为过 $\boldsymbol{p}$ 点且平行于 $A\boldsymbol{x}=\boldsymbol{0}$ 解集的一条直线. 见图 1-24.

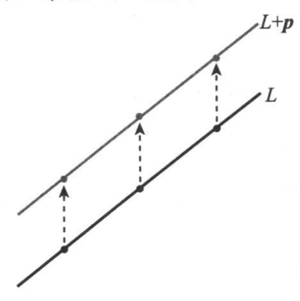


图 1-23 平移直线

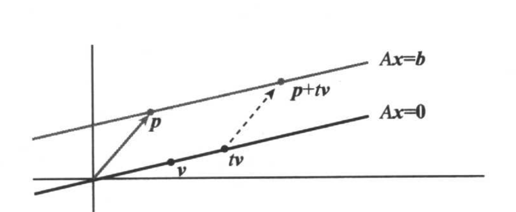


图 1-24 $A\boldsymbol{x}=\boldsymbol{b}$ 和 $A\boldsymbol{x}=\boldsymbol{0}$ 的平行解集

$A\boldsymbol{x}=\boldsymbol{b}$ 和 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集之间的关系如图 1-24 所示. 尽管含多个自由变量时，解集会比直线大，图 1-24 所示的关系仍可以推广到任意相容的方程 $A\boldsymbol{x}=\boldsymbol{b}$. 下面的定理给出了准确的陈述，定理的证明见习题 25.

【定理 6】 假设方程 $A\boldsymbol{x}=\boldsymbol{b}$ 对于某个给定的 $\boldsymbol{b}$ 相容，且令 $\boldsymbol{p}$ 为该方程的一个解. 则方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的解集是形如 $\boldsymbol{w}=\boldsymbol{p}+\boldsymbol{v}_h$ 的全体向量构成的集合，其中 $\boldsymbol{v}_h$ 是齐次方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的任意一个解.

定理 6 说明，如果 $A\boldsymbol{x}=\boldsymbol{b}$ 有解，则解集由 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集平移得到，并且可以使用 $A\boldsymbol{x}=\boldsymbol{b}$ 的任意特解 $\boldsymbol{p}$ 进行平移. 图 1-25 给出含两个自由变量的情况. 甚至当 $n>3$ 时，我们也可以想像得到相容方程组 $A\boldsymbol{x}=\boldsymbol{b}\,(\boldsymbol{b}\neq\boldsymbol{0})$ 的解集为一个非零单点或者不过原点的一条直线或一个平面.

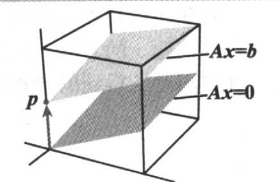


图 1-25 $A\boldsymbol{x}=\boldsymbol{b}$ 和 $A\boldsymbol{x}=\boldsymbol{0}$ 的平行解集

警告：定理 6 和图 1-25 仅适用于至少含有一个非零解的方程 $A\boldsymbol{x}=\boldsymbol{b}$. 若 $A\boldsymbol{x}=\boldsymbol{b}$ 无解，解集为空.

下面的运算法则概括了例题1、2和3中的计算.

将(相容方程组的)解集写成参数向量形式

1. 对增广矩阵进行行变换得到简化阶梯形式.
2. 用方程中出现的自由变量来表示每一个基本变量.
3. 把通解 $\mathbf{x}$ 写成向量形式，若存在自由变量，则 $\mathbf{x}$ 的元素与自由变量取值有关.
4. 把 $\mathbf{x}$ 分解成以自由变量为参数的向量（其中元素为数值）的线性组合形式.

基础练习

1. 下列每个方程都确定了 $\mathbf{R}^3$ 中的一个平面. 两个平面是否相交？如果是，请写出交集.

$$\begin{aligned} x_1+4x_2-5x_3&=0\\ 2x_1-x_2+8x_3&=9 \end{aligned}$$

54

2. 把 $10x_1-3x_2-2x_3=7$ 的通解写成参数向量形式. 使之与例题2中的解集联系起来.

习题 1.5

判断习题1~4中的方程组是否有非平凡解. 尽可能少地使用行变换.

1. $\begin{aligned} 2x_1-5x_2+8x_3&=0\\ -2x_1-7x_2+x_3&=0\\ 4x_1+2x_2+7x_3&=0 \end{aligned}$

2. $\begin{aligned} x_1-3x_2+7x_3&=0\\ -2x_1+x_2-4x_3&=0\\ x_1+2x_2+9x_3&=0 \end{aligned}$

3. $\begin{aligned} -3x_1+5x_2-7x_3&=0\\ -6x_1+7x_2+x_3&=0 \end{aligned}$

4. $\begin{aligned} -5x_1+7x_2+9x_3&=0\\ x_1-2x_2+6x_3&=0 \end{aligned}$

在习题5和6中，根据例题1和2的方法把已知的齐次方程组解集写成参数向量形式.

5. $\begin{aligned} x_1+3x_2+x_3&=0\\ -4x_1-9x_2+2x_3&=0\\ -3x_2-6x_3&=0 \end{aligned}$

6. $\begin{aligned} x_1+3x_2-5x_3&=0\\ x_1+4x_2-8x_3&=0\\ -3x_1-7x_2+9x_3&=0 \end{aligned}$

在习题7~12中，把 $A\mathbf{x}=\mathbf{0}$ 的所有解表示成参数向量形式，A 与已知矩阵行等价.

7. $\begin{bmatrix} 1 & 3 & -3 & 7\\ 0 & 1 & -4 & 5 \end{bmatrix}$

8. $\begin{bmatrix} 1 & -2 & -9 & 5\\ 0 & 1 & 2 & -6 \end{bmatrix}$

9. $\begin{bmatrix} 3 & -9 & 6\\ -1 & 3 & -2 \end{bmatrix}$

10. $\begin{bmatrix} 1 & 3 & 0 & -4\\ 2 & 6 & 0 & -8 \end{bmatrix}$

11. $\begin{bmatrix} 1 & -4 & -2 & 0 & 3 & -5\\ 0 & 0 & 1 & 0 & 0 & -1\\ 0 & 0 & 0 & 0 & 1 & -4\\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

12. $\begin{bmatrix} 1 & 5 & 2 & -6 & 9 & 0\\ 0 & 0 & 1 & -7 & 4 & -8\\ 0 & 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

13. 假设某个线性方程组的解集可以表示为 $x_1=5+4x_3$，$x_2=-2-7x_3$，x_3 为自由变量. 把解集用向量表示成 $\mathbf{R}^3$ 中的一条直线.
14. 假设某个线性方程组的解集可以表示为 $x_1=3x_4$，$x_2=8+x_4$，$x_3=2-5x_4$，x_4 为自由变量. 把解集用向量表示成 $\mathbf{R}^4$ 中的一条“直线”.
15. 根据例题3的方法，用参数向量形式表示下面方程组的解. 同样也给出解集的几何表示，并与习题5得到的解集进行比较.

$$\begin{aligned} x_1+3x_2+x_3&=1\\ -4x_1-9x_2+2x_3&=-1\\ -3x_2-6x_3&=-3 \end{aligned}$$

16. 如习题15，用参数向量形式表示下面方程组的解，并与习题6得到的解集进行几何比较.

$$\begin{aligned} x_1 + 3x_2 - 5x_3 &= 4 \\ x_1 + 4x_2 - 8x_3 &= 7 \\ -3x_1 - 7x_2 + 9x_3 &= -6 \end{aligned}$$

17. 求出并比较 $x_1 + 9x_2 - 4x_3 = 0$ 与 $x_1 + 9x_2 - 4x_3 = -2$ 的解集.
18. 求出并比较 $x_1 - 3x_2 + 5x_3 = 0$ 与 $x_1 - 3x_2 + 5x_3 = 4$ 的解集.

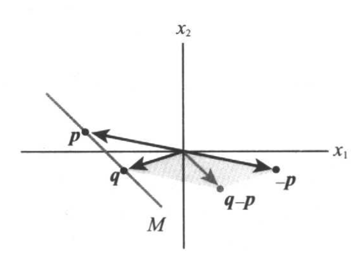
在习题 19 和 20 中，找出过 $\boldsymbol{a}$ 且平行于 $\boldsymbol{b}$ 的直线参数方程.

19. $\boldsymbol{a} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$　　20. $\boldsymbol{a} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} -7 \\ 8 \end{bmatrix}$

在习题 21 和 22 中，找出过 $\boldsymbol{p}$ 和 $\boldsymbol{q}$ 的直线 M 的参数方程. [提示：M 平行于向量 $\boldsymbol{q}-\boldsymbol{p}$. 见右图.]

21. $\boldsymbol{p} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$, $\boldsymbol{q} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$

22. $\boldsymbol{p} = \begin{bmatrix} -6 \\ 3 \end{bmatrix}$, $\boldsymbol{q} = \begin{bmatrix} 0 \\ -4 \end{bmatrix}$



习题 21、22 图　过 $\boldsymbol{p}$ 和 $\boldsymbol{q}$ 的直线

在习题 23 和 24 中，判断每个命题的真假，并说明理由.

23. a. 齐次方程总是相容的.
    b. 方程 $A\boldsymbol{x} = \boldsymbol{0}$ 给出了自身解集的显式表示.
    c. 齐次方程 $A\boldsymbol{x} = \boldsymbol{0}$ 有平凡解当且仅当方程至少含有一个自由变量.
    d. 方程 $\boldsymbol{x} = \boldsymbol{p} + t\boldsymbol{v}$ 表示过 $\boldsymbol{v}$ 且平行于 $\boldsymbol{p}$ 的一条直线.
    e. $A\boldsymbol{x} = \boldsymbol{b}$ 的解集是形如 $\boldsymbol{w} = \boldsymbol{p} + \boldsymbol{v}_h$ 的全体向量的集合，$\boldsymbol{v}_h$ 为方程 $A\boldsymbol{x} = \boldsymbol{0}$ 的任意解.

55

24. a. 如果 $\boldsymbol{x}$ 为 $A\boldsymbol{x} = \boldsymbol{0}$ 的一个非平凡解，则 $\boldsymbol{x}$ 中的每个元素都不为零.
    b. 方程 $\boldsymbol{x} = x_2\boldsymbol{u} + x_3\boldsymbol{v}$($\boldsymbol{u}$ 和 $\boldsymbol{v}$ 不互为倍数)表示过原点的一个平面.
    c. 当零向量为一个解时，$A\boldsymbol{x} = \boldsymbol{b}$ 是齐次方程.
    d. 把 $\boldsymbol{p}$ 加到一个向量上是：使向量沿着平行于 $\boldsymbol{p}$ 的方向移动.
    e. $A\boldsymbol{x} = \boldsymbol{b}$ 的解集是经过 $A\boldsymbol{x} = \boldsymbol{0}$ 的解集平移得到的.
25. 证明定理 6：
    a. 设 $\boldsymbol{p}$ 为 $A\boldsymbol{x} = \boldsymbol{b}$ 的一个解，则 $A\boldsymbol{p} = \boldsymbol{b}$. 设 $\boldsymbol{v}_h$ 为齐次方程 $A\boldsymbol{x} = \boldsymbol{0}$ 的任意解，且 $\boldsymbol{w} = \boldsymbol{p} + \boldsymbol{v}_h$. 证明 $\boldsymbol{w}$ 是 $A\boldsymbol{x} = \boldsymbol{b}$ 的一个解.
    b. 设 $\boldsymbol{w}$ 为方程 $A\boldsymbol{x} = \boldsymbol{b}$ 的一个解，定义 $\boldsymbol{v}_h = \boldsymbol{w} - \boldsymbol{p}$. 证明 $\boldsymbol{v}_h$ 是 $A\boldsymbol{x} = \boldsymbol{0}$ 的一个解. 这说明了 $A\boldsymbol{x} = \boldsymbol{b}$ 的每个解都有 $\boldsymbol{w} = \boldsymbol{p} + \boldsymbol{v}_h$ 的形式，$\boldsymbol{p}$ 为 $A\boldsymbol{x} = \boldsymbol{b}$ 的一个特解，$\boldsymbol{v}_h$ 为 $A\boldsymbol{x} = \boldsymbol{0}$ 的一个解.
26. 假设 $A\boldsymbol{x} = \boldsymbol{b}$ 有解，解释为什么当 $A\boldsymbol{x} = \boldsymbol{0}$ 只有平凡解时，$A\boldsymbol{x} = \boldsymbol{b}$ 的解恰好唯一.
27. 假设 A 为 3×3 零矩阵(所有元素均为 0). 求出方程 $A\boldsymbol{x} = \boldsymbol{0}$ 的解集.
28. 假设 $\boldsymbol{b} \neq \boldsymbol{0}$，那么 $A\boldsymbol{x} = \boldsymbol{b}$ 的解集是否是一个过原点的平面？给出解释.

在习题 29 ~ 32 中，(a)方程 $A\boldsymbol{x} = \boldsymbol{0}$ 是否有非平凡解？(b)方程 $A\boldsymbol{x} = \boldsymbol{b}$ 是否对于任意 $\boldsymbol{b}$ 至少有一个解？

29. A 为 3×3 矩阵，并且有三个主元位置.
30. A 为 3×3 矩阵，并且有两个主元位置.
31. A 为 3×2 矩阵，并且有两个主元位置.
32. A 为 2×4 矩阵，并且有两个主元位置.

33. 已知 $A=\begin{bmatrix}-2 & -6\\ 7 & 21\\ -3 & -9\end{bmatrix}$，通过观察找出 $A\boldsymbol{x}=\boldsymbol{0}$ 的一个非平凡解. [提示：考虑把方程 $A\boldsymbol{x}=\boldsymbol{0}$ 写成向量方程形式.]

34. 已知 $A=\begin{bmatrix}4 & -6\\ -8 & 12\\ 6 & -9\end{bmatrix}$，通过观察找出 $A\boldsymbol{x}=\boldsymbol{0}$ 的一个非平凡解.

35. 构造一个 3×3 的非零矩阵，使得向量 $\begin{bmatrix}1\\1\\1\end{bmatrix}$ 为 $A\boldsymbol{x}=\boldsymbol{0}$ 的一个解.

36. 构造一个 3×3 的非零矩阵，使得向量 $\begin{bmatrix}1\\-2\\1\end{bmatrix}$ 为 $A\boldsymbol{x}=\boldsymbol{0}$ 的一个解.

37. 构造一个 2×2 矩阵 A，使得 $A\boldsymbol{x}=\boldsymbol{0}$ 的解集为 $\mathbb{R}^2$ 中过 $(4,1)$ 和原点的直线. 然后找出 $\mathbb{R}^2$ 中的向量 $\boldsymbol{b}$，使得 $A\boldsymbol{x}=\boldsymbol{b}$ 的解集不是 $\mathbb{R}^2$ 中平行于 $A\boldsymbol{x}=\boldsymbol{0}$ 解集的直线.

38. 设 A 为 3×3 矩阵，$\boldsymbol{y}$ 是 $\mathbb{R}^2$ 中向量，且 $A\boldsymbol{x}=\boldsymbol{y}$ 无解. 讨论是否存在 $\mathbb{R}^3$ 中向量 $\boldsymbol{z}$，使得方程 $A\boldsymbol{x}=\boldsymbol{z}$ 有唯一解?

39. 设 A 为 $m\times n$ 矩阵，$\boldsymbol{u}$ 为 $\mathbb{R}^n$ 中满足方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的向量. 请说明对于任意的数量 c，$c\boldsymbol{u}$ 都满足 $A\boldsymbol{x}=\boldsymbol{0}$. [即是说明 $A(c\boldsymbol{u})=\boldsymbol{0}$.]

40. 设 A 为 $m\times n$ 矩阵，$\boldsymbol{u}$ 和 $\boldsymbol{v}$ 为 $\mathbb{R}^n$ 中的向量，满足 $A\boldsymbol{u}=\boldsymbol{0}$ 和 $A\boldsymbol{v}=\boldsymbol{0}$. 为什么 $A(\boldsymbol{u}+\boldsymbol{v})$ 一定为零向量? 然后解释为什么对任意两个数量 c 和 d，都有 $A(c\boldsymbol{u}+d\boldsymbol{v})=\boldsymbol{0}$.

基础练习答案

56

1. 对增广矩阵进行行化简：

$$\begin{bmatrix}1 & 4 & -5 & 0\\ 2 & -1 & 8 & 9\end{bmatrix}\sim\begin{bmatrix}1 & 4 & -5 & 0\\ 0 & -9 & 18 & 9\end{bmatrix}\sim\begin{bmatrix}1 & 0 & 3 & 4\\ 0 & 1 & -2 & -1\end{bmatrix},$$

$$\begin{aligned} x_1 \quad +3x_3 &= 4\\ x_2 - 2x_3 &= -1 \end{aligned}$$

因此 $x_1=4-3x_3$，$x_2=-1+2x_3$，x_3 为自由变量. 参数向量形式的通解为

$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}4-3x_3\\-1+2x_3\\x_3\end{bmatrix}=\underbrace{\begin{bmatrix}4\\-1\\0\end{bmatrix}}_{\boldsymbol{p}}+x_3\underbrace{\begin{bmatrix}-3\\2\\1\end{bmatrix}}_{\boldsymbol{v}}$$

两个平面的交集是一条过点 $\boldsymbol{p}$ 且沿着方向 $\boldsymbol{v}$ 的直线.

2. 增广矩阵 $[10 \quad -3 \quad -2 \quad 7]$ 与 $[1 \quad -0.3 \quad -0.2 \quad 0.7]$ 行等价，通解为 $x_1=0.7+0.3x_2+0.2x_3$，x_2 和 x_3 是自由变量. 即

$$\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0.7+0.3x_2+0.2x_3\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0.7\\0\\0\end{bmatrix}+x_2\begin{bmatrix}0.3\\1\\0\end{bmatrix}+x_3\begin{bmatrix}0.2\\0\\1\end{bmatrix}=\boldsymbol{p}+x_2\boldsymbol{u}+x_3\boldsymbol{v}$$

非齐次方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的解集是一个经过平移的平面 $\boldsymbol{p}+\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$. 该平面过点 $\boldsymbol{p}$，且平行于例题 2 中齐次方程的解集.

1.6 线性方程组的应用

你可能希望一个与线性代数有关的实际问题有唯一解，或者无解. 本节的目的是要说明在应用中，有多个解的线性方程组是怎样产生的. 这里所说的应用来源于经济学、化学和网络流.

1.6.1 经济学中的齐次方程组

本章实例介绍中提到的含500个未知量的500个方程，现在被称为列昂惕夫“投入－产出”模型.[1] 2.6节将详细讨论这个模型，那时我们的理论和记号都很完备. 不过，我们当前只考虑列昂惕夫提出的一种更简单的“交换模型”.

假设一个国家的经济分为很多行业，例如制造业、通讯业、娱乐业和服务行业等. 我们知道每个部门一年的总产出，并准确了解其产出如何在经济的其他部门之间分配或“交易”. 把一个部门产出的总货币价值称为该产出的**价格**(price). 列昂惕夫证明了如下结论.

57

存在赋给各部门总产出的平衡价格，使得每个部门的投入与产出都相等.

下面的例题说明如何求平衡价格.

【例题1】 假设一个经济系统由煤炭、电力、钢铁行业组成，每个行业的产出在各个行业中的分配如表1-1所示，每一列中的元素表示占该行业总产出的比例.

以表1-1的第2列为例，电力行业的总产出分配如下：40%分配到煤炭行业，50%分配到钢铁行业，余下的10%分配到电力行业(电力行业把这10%当作部门营运所需的投入). 因为考虑了所有的产出，所以每一列的小数加起来必须等于1.

把煤炭、电力、钢铁行业每年总产出的价格(即货币价值)分别用p_C，p_E和p_S表示. 试求使得每个行业的投入与产出都相等的平衡价格.

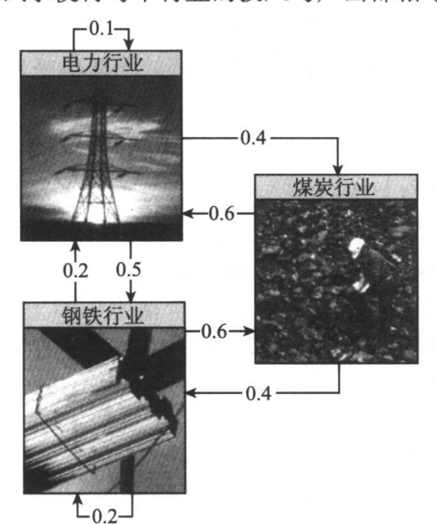


表1-1 一个简单的经济系统

产出分配			购买者
煤炭	电力	钢铁	
0.0	0.4	0.6	煤炭
0.6	0.1	0.2	电力
0.4	0.5	0.2	钢铁

1. 见 Wassily W. Leontief, “Input-Output Economics”, *Scientific American*, October 1951, pp. 15－21.

解：表1-1中，沿列可以看出每个行业的产出分配到何处，沿行则可以看出这个行业所需的投入. 例如，第1行说明煤炭行业接收(购买)了40%的电力产出和60%的钢铁产出，由于电力和钢铁的总产出价格分别是p_E和p_S，因此煤炭行业必须分别向电力行业和钢铁行业支付$0.4p_E$和$0.6p_S$美元. 煤炭行业的总支出为$0.4p_E+0.6p_S$. 为了使煤炭行业的收入p_C等于它的支出，我们希望

$$p_C=0.4p_E+0.6p_S \tag{1}$$

交易表格的第2行说明了电力行业分别要向煤炭、电力、钢铁各行业支付

58

$0.6p_C$，$0.1p_E$和$0.2p_S$. 因此电力行业的收支平衡条件为

$$p_E=0.6p_C+0.1p_E+0.2p_S \tag{2}$$

最后，交易表格的第3行导出了最后一个收支平衡条件：

$$p_S=0.4p_C+0.5p_E+0.2p_S \tag{3}$$

为求解由方程(1)、(2)和(3)构成的方程组，将所有的未知量移到方程左边，然后合并同类项. [例如，(2)式的左边把$p_E-0.1p_E$写成$0.9p_E$]

$$\begin{aligned} p_C-0.4p_E-0.6p_S&=0\\ -0.6p_C+0.9p_E-0.2p_S&=0\\ -0.4p_C-0.5p_E+0.8p_S&=0 \end{aligned}$$

接下来进行行化简，为了简单起见，保留两位小数.

$$\begin{bmatrix} 1 & -0.4 & -0.6 & 0\\ -0.6 & 0.9 & -0.2 & 0\\ -0.4 & -0.5 & 0.8 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -0.4 & -0.6 & 0\\ 0 & 0.66 & -0.56 & 0\\ 0 & -0.66 & 0.56 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -0.4 & -0.6 & 0\\ 0 & 0.66 & -0.56 & 0\\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -0.4 & -0.6 & 0\\ 0 & 1 & -0.85 & 0\\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -0.94 & 0\\ 0 & 1 & -0.85 & 0\\ 0 & 0 & 0 & 0 \end{bmatrix}$$

通解为$p_C=0.94p_S$，$p_E=0.85p_S$，其中p_S为自由变量. 经济系统的平衡价格向量有如下形式

$$\boldsymbol{p}=\begin{bmatrix} p_C\\ p_E\\ p_S \end{bmatrix}=\begin{bmatrix} 0.94p_S\\ 0.85p_S\\ p_S \end{bmatrix}=p_S\begin{bmatrix} 0.94\\ 0.85\\ 1 \end{bmatrix}$$

每个p_S的(非负)取值都确定一个平衡价格的取值. 例如，我们取p_S为100百万美元(或者说1亿美元)，则$p_C=94$，$p_E=85$. 即如果煤炭行业产出价格为0.94亿美元，则电力行业产出价格为0.85亿美元，钢铁行业产出价格为1亿美元，那么每个行业的收入和支出相等. ■

1.6.2　配平化学方程式

化学方程式表示化学反应中消耗和产生的物质的量. 例如，燃烧丙烷时丙烷(C_3H_8)和氧气(O_2)结合，生成二氧化碳(CO_2)和水(H_2O)，其方程式为：

$$(x_1)C_3H_8+(x_2)O_2\rightarrow(x_3)CO_2+(x_4)H_2O \tag{4}$$

为了“配平”这个方程式，化学家必须找出一列$x_1,\cdots,x_4$，使得方程式左端的碳原子(C)、氢原子(H)和氧原子(O)的总数与右端对应的原子总数相等(因为化学反应

中原有的原子不可能消失，也不可能产生新原子）.

59

配平化学方程式的一个系统的方法，就是建立能描述反应过程中每种原子数目的向量方程. 方程(4)包含了3种不同的原子(碳、氢、氧)，于是在$\mathbf{R}^3$中为(4)的每一种反应物和生成物构造如下向量，在其中列出每个分子所包含的不同原子的数目：

$$C_3H_8:\begin{bmatrix}3\\8\\0\end{bmatrix},\quad O_2:\begin{bmatrix}0\\0\\2\end{bmatrix},\quad CO_2:\begin{bmatrix}1\\0\\2\end{bmatrix},\quad H_2O:\begin{bmatrix}0\\2\\1\end{bmatrix}\begin{matrix}\leftarrow\text{碳}\\\leftarrow\text{氢}\\\leftarrow\text{氧}\end{matrix}$$

为了配平方程式(4)，系数$x_1,\cdots,x_4$必须满足

$$x_1\begin{bmatrix}3\\8\\0\end{bmatrix}+x_2\begin{bmatrix}0\\0\\2\end{bmatrix}=x_3\begin{bmatrix}1\\0\\2\end{bmatrix}+x_4\begin{bmatrix}0\\2\\1\end{bmatrix}$$

为了求解，把所有的项都移到左端(注意改变第3个和第4个向量的符号)：

$$x_1\begin{bmatrix}3\\8\\0\end{bmatrix}+x_2\begin{bmatrix}0\\0\\2\end{bmatrix}+x_3\begin{bmatrix}-1\\0\\-2\end{bmatrix}+x_4\begin{bmatrix}0\\-2\\-1\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

对该方程的增广矩阵进行行化简，得到通解

$$x_1=\frac{1}{4}x_4,x_2=\frac{5}{4}x_4,x_3=\frac{3}{4}x_4,\quad x_4\text{是自由变量}$$

由于化学方程式中的系数必须为整数，取$x_4=4$，此时$x_1=1$，$x_2=5$且$x_3=3$. 配平后的方程式为

$$C_3H_8+5O_2\rightarrow 3CO_2+4H_2O$$

如果将每个系数翻倍，方程式仍然平衡. 不过，在大多数场合下化学家更倾向于使用尽可能小的整数来配平方程式.

1.6.3 网络流

当科学家、工程师和经济学家研究网络中的流量时，线性方程组就自然产生了. 例如，城市规划者和交通工程师监控城市道路网格内的交通流量，电气工程师计算电路中流经的电流，经济学家分析产品通过批发商和零售商网络从生产者到消费者的分配. 大多数网络问题中的方程组包含了数百甚至上千个变量和方程.

一个网络由一个点集以及连接部分或全部点的直线或弧线构成. 网络中的点称作联结点(或节点)，网络中的连接线称作分支. 每一分支中的流量方向已经指定，并且流量(或流速)已知或者已标为变量.

60

网络流的基本设想是网络中流入和流出的总量相等，并且每个联结点流入和流出的总量也相等. 例如，图1-26说明了30单位的流量从一个分支流入联结点，x_1和x_2表示该联结点从其他分支流出的流量. 因为流量在每个联结点守恒，所以我们一定有$x_1+x_2=30$. 在类似的网络模式中，每个联结点的流量都可以用一个线性方程组来表示. 网络分析要解决的问题就是：在部分信息(如网络的输入量)已知的情况下，确定每一分支中的流量.

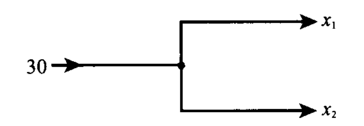


图1-26 一个联结点，或结点

【例题 2】 图 1-27 中的网络给出了在下午一两点钟，巴尔的摩市区一些单行道的交通流量(以每小时的汽车数量来度量)．试确定网络的流量模式.

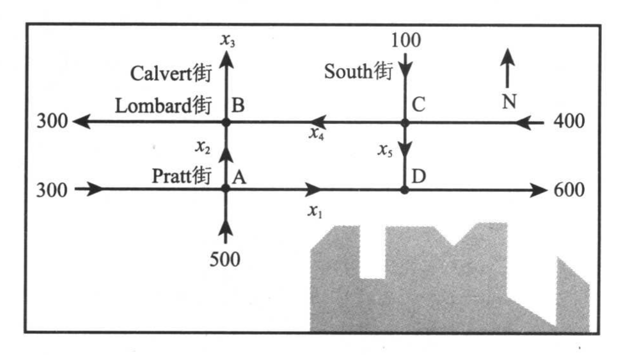


图 1-27 巴尔的摩街道

解： 写出表示流量的方程，然后求出方程组的通解．首先标出街道交叉点(联结点)和分支中未知的流量，如图 1-27 所示．设定每个交叉点的流入和流出相等.

交叉点	流入		流出
A	$300+500$	=	x_1+x_2
B	x_2+x_4	=	$300+x_3$
C	$100+400$	=	x_4+x_5
D	x_1+x_5	=	600

61

此外，该网络的总流入$(500+300+100+400)$等于网络的总流出$(300+x_3+600)$，化简得 $x_3=400$．把这个方程与整理后的前四个方程联立起来，得到如下方程组：

$$\begin{aligned}
x_1+x_2\qquad\qquad\quad &=800\\
x_2-x_3+x_4\quad &=300\\
x_4+x_5&=500\\
x_1\qquad\qquad\qquad+x_5&=600\\
x_3\qquad\qquad &=400
\end{aligned}$$

对应的增广矩阵行化简得到

$$\begin{aligned}
x_1\qquad\quad +x_5&=600\\
x_2\qquad -x_5&=200\\
x_3\qquad\quad &=400\\
x_4+x_5&=500
\end{aligned}$$

网络的流量模式表示为

$$\begin{cases}
x_1=600-x_5\\
x_2=200+x_5\\
x_3=400\\
x_4=500-x_5\\
x_5\text{ 是自由变量}
\end{cases}$$

网络分支中的负流量表示与模型中指定的方向相反. 由于街道是单行道，因此变量不能取负值. 这就导致变量在取正值时也有一定的局限. 例如，由于 x_4 不能为负，所以 $x_5 \leqslant 500$. 变量的其他约束将在基础练习 2 中讨论. ■

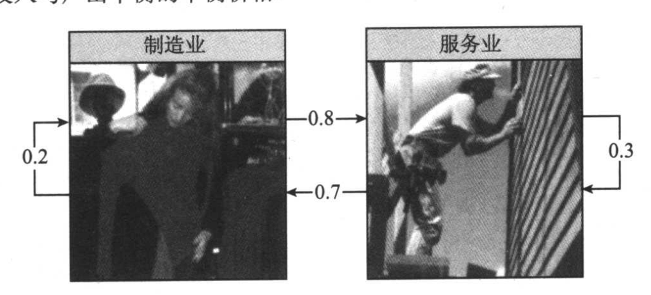
基础练习

1. 假设一个经济系统由 3 个行业组成，农业、矿业和制造业. 农业分别出售 5% 和 30% 的产品给矿业和制造业，保留余下部分. 矿业分别出售 20% 和 70% 的产品给农业和制造业，保留余下部分. 制造业分别出售 20% 和 30% 的产品给农业和矿业，保留余下部分. 试确定这个经济系统的交易表格，要求用表格中的列描述每个行业的产出在三个行业之间的分配.
2. 考虑例题 2 中的网络流，确定 x_1 和 x_2 的取值范围. [提示：例题 2 中得到了 $x_5 \leqslant 500$. 对于 x_1 和 x_2 这意味着什么？另外再利用 $x_5 \geqslant 0$，可以得到什么结论？]

62

习题 1.6

1. 假设一个经济系统只有两个部门：制造业、服务业. 每一年，制造业卖出 80% 产品给服务业，保留余下部分；服务业卖出 70% 产品给制造业，保留余下部分. 找出使得制造业和服务业每年的投入与产出平衡的平衡价格.



习题 1 图

2. 假设用日元替换美元来衡量各个行业的产出价格，求出例题 1 中经济系统的另一个平衡价格. 这样是否在某种方式上改变了这个问题？
3. 考虑一个含有 3 个行业的经济系统：五金化工、燃料和电力、机械. 化工行业分别卖出 30% 和 50% 的产品给电力行业和机械行业，保留余下部分. 电力行业分别卖出 80% 和 10% 的产品给化工行业和机械行业，保留余下部分. 机械行业各卖出 40% 的产品给化工行业和电力行业，保留余下部分.
   a. 构造这个经济系统的交易表格.
   b. 改进方程组以便得到使每个行业投入等于产出的价格. 然后写出可以通过行化简得到这些价格的增广矩阵.
   c. [M]机械行业的产出为 100 单位时，求出平衡价格.
4. 假设一个经济系统有四个行业：农业(A)、能源(E)、制造(M)和通信(T). A 行业分别卖出 10% 和 25% 的产品给 E 行业和 M 行业，保留余下部分. E 行业分别卖出 30%、35% 和 25% 的产品给 A 行业、M 行业和 T 行业，保留余下部分. M 行业分别卖出 30%、15% 和 40% 的产品给 A 行业、E 行业和 T 行业，保留余下部分. T 行业分别卖出 20%、10% 和 30% 的产品给 A 行业、E 行业和 M 行业，保留余下部分.
   a. 构造这个经济系统的交易表格.

b. [**M**]求出这个经济系统的平衡价格.

用本节中讨论的向量方程方法来配平习题5~10中的化学方程式.

5. 硼硫化合物与水发生剧烈的反应，生成硼酸和硫化氢气体(臭鸡蛋味道). 待平衡的方程式为

$$B_2S_3 + H_2O \to H_3BO_3 + H_2S$$

[对每一种化合物，构造一个列出硼、硫、氢和氧原子数目的向量.]

6. 当磷酸钠溶液和硝酸钡溶液混合后，产生磷酸钡和硝酸钠. 待平衡的方程式为

$$Na_3PO_4 + Ba(NO_3)_2 \to Ba_3(PO_4)_2 + NaNO_3$$

[对每一种化合物，构造一个列出钠、磷、氧、钡和氮原子数目的向量. 例如，硝酸钡对应于(0,0,6,1,2).]

7. Alka苏打水含有碳酸氢钠($NaHCO_3$)和柠檬酸($H_3C_6H_5O_7$). 药片在水中溶解，按照如下反应生成柠檬酸钠、水和二氧化碳(气体).

$$NaHCO_3 + H_3C_6H_5O_7 \to Na_3C_6H_5O_7 + H_2O + CO_2$$

8. 如下高锰酸钾($KMnO_4$)与硫酸锰在水中的反应生成二氧化锰、硫酸钾和硫酸.

$$KMnO_4 + MnSO_4 + H_2O \to MnO_2 + K_2SO_4 + H_2SO_4$$

[对每一种化合物，构造一个列出钾(K)、锰、氧、硫和氢原子数目的向量.]

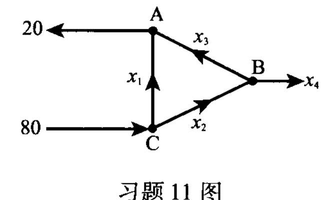
9. [**M**]如果可能，在配平化学反应式的计算过程中使用有理分式，不进行舍入.

$$PbN_6 + CrMn_2O_8 \to Pb_3O_4 + Cr_2O_3 + MnO_2 + NO$$

10. [**M**]如下的化学反应式可以在工业中应用，如砷化三氢(AsH_3)的生产. 配平这一方程式，注意在计算过程中使用有理分式，不进行舍入.

$$MnS + As_2Cr_{10}O_{35} + H_2SO_4 \to HMnO_4 + AsH_3 + CrS_3O_{12} + H_2O$$

11. 给出如右图所示的流量模式. 假设所有的流量都非负，x_3的最大可能值是多少?



习题11图

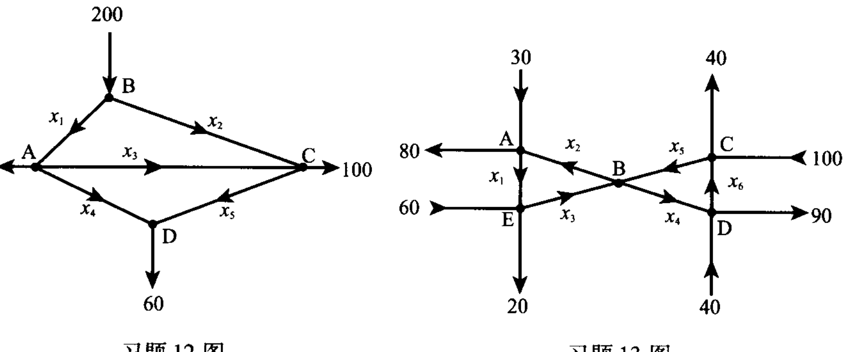
12. a. 给出如下图所示的高速公路网络的流量模式(速度为：车数/分).

    b. 当流量为x_4的路面关闭时，表示出流量模式.

    c. 流量$x_4=0$时，x_1的最小值是多少?

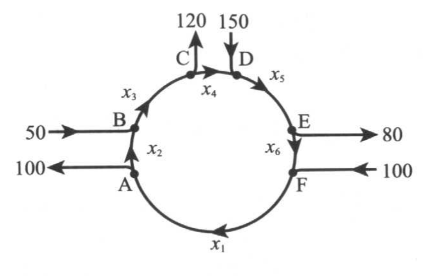
13. a. 给出如下图所示的流量模式.

    b. 假设流量必须沿着指定的方向，找出标记为x_2，x_3，x_4和x_5的分支中最小流量.



习题12图　　习题13图

14. 英国的道路交叉处通常建成单行的小环岛，如下图. 假设交通行进方向必须如图示那样，请求出该网络流的通解. 找出x_6的最小可能值.



习题 14 图

基础练习答案

1. 因为必须考虑所有的产出，把百分比写成小数形式. 则每一列的和都应是 1.

产出分配			购买者
农业	矿业	制造业	
0.65	0.20	0.20	农业
0.05	0.10	0.30	矿业
0.30	0.70	0.50	制造业

64

2. 因为 $x_1 \leqslant 500$，从 x_1 和 x_2 的方程得到 $x_1 \geqslant 100$，$x_2 \leqslant 700$. $x_5 \geqslant 0$ 蕴含了 $x_1 \leqslant 600$ 和 $x_2 \geqslant 200$. 因此有 $100 \leqslant x_1 \leqslant 600$，$200 \leqslant x_2 \leqslant 700$.

1.7 线性无关

我们可以从另一个角度来研究 1.5 节中的齐次方程，即把它们写成向量方程的形式. 这样就把注意力从 $A\boldsymbol{x}=\boldsymbol{0}$ 的未知解转移到向量方程中的向量上了.

例如，考虑如下方程

$$x_1\begin{bmatrix}1\\2\\3\end{bmatrix}+x_2\begin{bmatrix}4\\5\\6\end{bmatrix}+x_3\begin{bmatrix}2\\1\\0\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix} \tag{1}$$

$x_1=x_2=x_3=0$ 时，方程显然有一个平凡解. 与 1.5 节相同，主要的问题是平凡解是否是唯一解.

【定义】 如果方程 $x_1\boldsymbol{v}_1+x_2\boldsymbol{v}_2+\cdots+x_p\boldsymbol{v}_p=\boldsymbol{0}$ 只有平凡解，则称 $\mathbf{R}^n$ 中带指标的向量集 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$**线性无关**(linearly independent). 若存在不全为零的权重 $c_1,\cdots,c_p$，使得

$$c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_p\boldsymbol{v}_p=\boldsymbol{0} \tag{2}$$

则称集合 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$**线性相关**(linearly dependent).

权重不全为零时，方程(2)称为 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 的**线性相关关系**(linearly dependent relation). 一个向量集线性相关当且仅当它不是线性无关的. 简单地说，我们说 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 线性相关，指的是 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是线性相关集. 类似地，还可以定义线性无关集.

【例题 1】 设 $\boldsymbol{v}_1=\begin{bmatrix}1\\2\\3\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}4\\5\\6\end{bmatrix}$，$\boldsymbol{v}_3=\begin{bmatrix}2\\1\\0\end{bmatrix}$.

a. 判断集合$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$是否线性无关.

65

b. 试求$\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$间的线性相关关系.

解：

a. 我们必须判断上面的方程(1)是否有非平凡解. 对相应的增广矩阵进行行变换得到

$$\begin{bmatrix}1&4&2&0\\2&5&1&0\\3&6&0&0\end{bmatrix}\sim\begin{bmatrix}1&4&2&0\\0&-3&-3&0\\0&0&0&0\end{bmatrix}$$

显然，x_1和x_2为基本变量，x_3是自由变量. x_3的每个非零的取值都确定了(1)的一个非平凡解. 因此$\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$是线性相关的.

b. 要求$\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$间的线性相关关系，需要对增广矩阵进行完全的行化简，得到的新方程组为：

$$\begin{bmatrix}1&0&-2&0\\0&1&1&0\\0&0&0&0\end{bmatrix}\quad\begin{aligned}x_1\quad-2x_3&=0\\x_2+x_3&=0\\0&=0\end{aligned}$$

因此$x_1=2x_3$，$x_2=-x_3$，其中x_3是自由变量. x_3任取一个非零值，比如$x_3=5$，则有$x_1=10$，$x_2=-5$. 代入这些值到(1)中得到

$$10\boldsymbol{v}_1-5\boldsymbol{v}_2+5\boldsymbol{v}_3=\boldsymbol{0}$$

这就是$\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$间的一种(共有无穷多种)可能的线性相关关系. ■

1.7.1 矩阵列的线性无关

假设给定了矩阵$A=[\boldsymbol{a}_1\quad\boldsymbol{a}_2\cdots\boldsymbol{a}_n]$而不是一个向量集合. 矩阵方程$A\boldsymbol{x}=\boldsymbol{0}$可以写成

$$x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\cdots+x_n\boldsymbol{a}_n=\boldsymbol{0}$$

A中列之间的线性相关关系对应于$A\boldsymbol{x}=\boldsymbol{0}$的一个非平凡解. 因此我们有下述重要结论.

矩阵A中的列线性无关当且仅当方程$A\boldsymbol{x}=\boldsymbol{0}$只有平凡解.　　(3)

66

【例题2】 判断矩阵$A=\begin{bmatrix}0&1&4\\1&2&-1\\5&8&0\end{bmatrix}$中的列是否线性无关.

解： 考虑$A\boldsymbol{x}=\boldsymbol{0}$，对增广矩阵进行行化简：

$$\begin{bmatrix}0&1&4&0\\1&2&-1&0\\5&8&0&0\end{bmatrix}\sim\begin{bmatrix}1&2&-1&0\\0&1&4&0\\0&-2&5&0\end{bmatrix}\sim\begin{bmatrix}1&2&-1&0\\0&1&4&0\\0&0&13&0\end{bmatrix}$$

容易看出方程组含3个基本变量，且不含自由变量. 因此方程$A\boldsymbol{x}=\boldsymbol{0}$只有平凡解，于是$A$中的列线性无关. ■

1.7.2 含一个或两个向量的集合

只含有一个向量$\boldsymbol{v}$的集合是线性无关的当且仅当$\boldsymbol{v}$是非零向量，这是因为当$\boldsymbol{v}\neq\boldsymbol{0}$时，向量方程$x_1\boldsymbol{v}=\boldsymbol{0}$只有平凡解. 零向量是线性相关的，因为$x_1\boldsymbol{0}=\boldsymbol{0}$有多个非平

凡解.

下面的例题将揭示仅含两个向量的线性相关集的本质.

【例题 3】 判断下面的向量集合是否线性无关.

$$\text{a. } \boldsymbol{v}_1=\begin{bmatrix}3\\1\end{bmatrix},\ \boldsymbol{v}_2=\begin{bmatrix}6\\2\end{bmatrix}\quad \text{b. } \boldsymbol{v}_1=\begin{bmatrix}3\\2\end{bmatrix},\ \boldsymbol{v}_2=\begin{bmatrix}6\\2\end{bmatrix}$$

解:

a. 注意 $\boldsymbol{v}_2$是 $\boldsymbol{v}_1$的倍数，即 $\boldsymbol{v}_2=2\boldsymbol{v}_1$，因此 $-2\boldsymbol{v}_1+\boldsymbol{v}_2=\boldsymbol{0}$，说明$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$线性相关.

b. 向量 $\boldsymbol{v}_1$和 $\boldsymbol{v}_2$之间没有倍数关系，它们是否线性相关？假设 c 和 d 满足

$$c\boldsymbol{v}_1+d\boldsymbol{v}_2=\boldsymbol{0}$$

如果 $c\neq 0$，则我们可以用 $\boldsymbol{v}_2$表示出 $\boldsymbol{v}_1$，即 $\boldsymbol{v}_1=(-d/c)\boldsymbol{v}_2$. 因为 $\boldsymbol{v}_1$不是 $\boldsymbol{v}_2$的倍数，所以这个结果不可能成立. 因此 c 必定为 0. 类似地，d 也必定为 0. 因此$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$是线性无关集. ■

例题 3 中的讨论说明，无需进行行变换，可以仅通过观察来判断含有两个向量的集合是否线性相关，即只需检查是否其中一个向量是另一个向量的数量倍数(该判别法只适用于仅含两个向量的集合).

向量集合$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$中如果有一个向量是另一个向量的数量倍数，则该集合线性相关. 该集合线性无关当且仅当两个向量之间没有倍数关系.

用几何语言表述，即，两个向量线性相关当且仅当它们位于同一条过原点的直线上. 图 1-28 绘出了例题 3 中的两个向量.

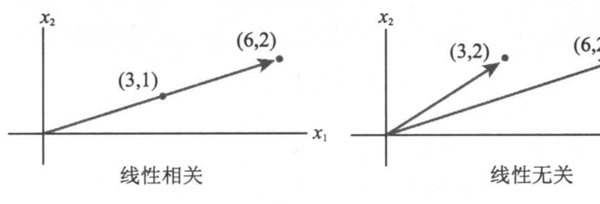


图 1-28

1.7.3 含两个或两个以上向量的集合

下面定理的证明与求解例题 3 类似，证明细节在本节末尾给出.

【定理 7】 线性相关集的特征

一个含两个或两个以上向量的集合 $S=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$线性相关当且仅当 S 中至少有一个向量是其余向量的线性组合. 事实上，如果 S 线性相关，且 $\boldsymbol{v}_1\neq\boldsymbol{0}$，则存在向量 $\boldsymbol{v}_j(j>1)$是它前面向量 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{j-1}$的线性组合.

警告：定理 7 并没有说线性相关集中的每个向量都是它前面向量的线性组合. 线性相关集中可能存在某个向量不是其余向量的线性组合. 见基础练习 3.

【例题 4】 设 $\boldsymbol{u}=\begin{bmatrix}3\\1\\0\end{bmatrix}$, $\boldsymbol{v}=\begin{bmatrix}1\\6\\0\end{bmatrix}$. 表示出 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 张成的集合，并解释为什么向量

$\boldsymbol{w} \in \mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$当且仅当$\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$线性相关.

解: 向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 之间没有倍数关系，因此它们线性无关，进而张成 $\mathbf{R}^3$ 的一个平面(见 1.3 节). 事实上，$\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$就是 x_1x_2 - 平面($x_3=0$). 如果 $\boldsymbol{w}$ 是 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的线性组合，则根据定理 7 有 $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$线性相关. 反过来，假设$\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$线性相关，根据定理 7，则存在$\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$中的一个向量，该向量是它前面所有向量的线性组合(因为 $\boldsymbol{u}\neq\boldsymbol{0}$). 因为 $\boldsymbol{u}$ 不是 $\boldsymbol{v}$ 的倍数，这个向量一定是 $\boldsymbol{w}$，故 $\boldsymbol{w}\in\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$. 见图 1-29.

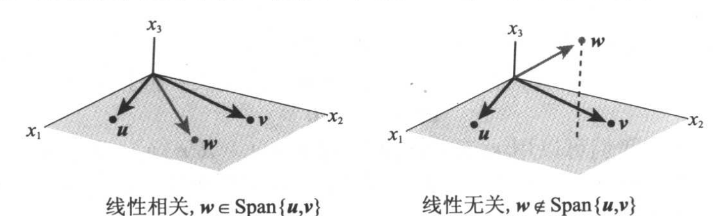


图 1-29　$\mathbf{R}^3$ 中的线性相关

■

例题 4 的结论可以推广到任意 $\mathbf{R}^3$ 中的集合$\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$(其中 $\boldsymbol{u}$, $\boldsymbol{v}$ 线性无关)上：集合$\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$线性相关当且仅当 $\boldsymbol{w}$ 属于 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 张成的平面.

接下来的两个定理描述了一个集合必定线性相关的特殊情况. 定理 8 也是后面章节中要频繁使用的一个关键结论.

68

【定理 8】　如果一个集合含有的向量个数多于每个向量中所含元素的个数，则该集合线性相关. 即，如果 $p>n$，则任意 $\mathbf{R}^n$ 中的集合$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$线性相关.

证明: 设 $A=[\boldsymbol{v}_1\ \ \cdots\ \ \boldsymbol{v}_p]$. 则 A 为 $n\times p$ 阶矩阵，并且方程 $A\boldsymbol{x}=\boldsymbol{0}$ 对应于含有 p 个未知量 n 个方程的方程组. 如果 $p>n$，则变量个数多于方程的个数. 因此 $A\boldsymbol{x}=\boldsymbol{0}$ 有非平凡解，并且 A 中的列线性相关. 图 1-30 用矩阵的形式重述了这个定理.

$$n\overset{p}{\begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}}$$

图 1-30　如果 $p>n$，则列向量线性相关

■

警告：定理 8 并没有考虑集合中向量个数不大于每个向量中所含元素的个数这种情况.

【例题 5】　根据定理8，向量$\begin{bmatrix}2\\1\end{bmatrix}$, $\begin{bmatrix}4\\-1\end{bmatrix}$, $\begin{bmatrix}-2\\2\end{bmatrix}$线性相关. 因为集合中有 3 个向量，但每个向量中仅有 2 个元素. 注意，尽管它们线性相关，但是其中任意向量都不是其他向量的倍数. 见图1-31.

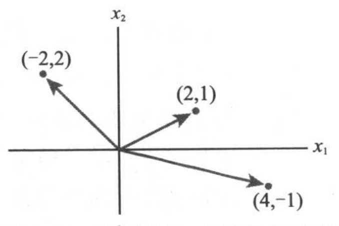


图 1-31　$\mathbf{R}^2$ 中的一个线性相关集

■

【定理 9】　如果 $\mathbf{R}^n$ 中的一个集合 $S=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$包含零向量，则该集合线性相关.

证明: 对向量重新编号，可以假设 $\boldsymbol{v}_1=\boldsymbol{0}$. 则方程 $1\boldsymbol{v}_1+0\boldsymbol{v}_2+\cdots+0\boldsymbol{v}_p=\boldsymbol{0}$ 表明 S 线性相关.

■

【例题 6】　通过观察判断给出的集合是否线性相关.

a. $\begin{bmatrix}1\\7\\6\end{bmatrix}, \begin{bmatrix}2\\0\\9\end{bmatrix}, \begin{bmatrix}3\\1\\5\end{bmatrix}, \begin{bmatrix}4\\1\\8\end{bmatrix}$ b. $\begin{bmatrix}2\\3\\5\end{bmatrix}, \begin{bmatrix}0\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\8\end{bmatrix}$ c. $\begin{bmatrix}-2\\4\\6\\10\end{bmatrix}, \begin{bmatrix}3\\-6\\-9\\15\end{bmatrix}$

解:

a. 该集合含有 4 个向量，每个向量仅有 3 个元素. 根据定理 8 知该集合线性相关.

b. 因为向量个数不多于每个向量中所含元素的个数，此处不能应用定理 8. 但集合中有零向量，根据定理 9 知该集合线性相关.

c. 比较两个向量中的对应元素，第二个向量似乎是第一个的 $-3/2$ 倍. 这个关系对向量的前三对元素成立，但对第四对元素并不成立. 因此它们之间没有倍数关系，两者线性无关. ■

69

一般来说，需要反复阅读这一节才能理解线性无关这个重要概念. 学习指南中关于本节的注解将帮助你对线性代数的关键思想形成直观映像. 例如，下面的证明演示了如何应用线性无关的定义，值得认真阅读.

定理 7（线性相关集的特征）的证明： 如果存在 $\mathbf{v}_j \in S$ 是其余向量的线性组合，则等式两端可以同时减去 $\mathbf{v}_j$，得到一个线性相关关系，其中 $\mathbf{v}_j$ 的权重（-1）非零. 例如，若 $\mathbf{v}_1 = c_2\mathbf{v}_2 + c_3\mathbf{v}_3$，则 $\mathbf{0} = (-1)\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + 0\mathbf{v}_4 + \cdots + 0\mathbf{v}_p$. 因此 S 是线性相关的.

反过来，假设 S 线性相关. 如果 $\mathbf{v}_1$ 为 $\mathbf{0}$，则它是 S 中其余向量的一个（平凡）线性组合. 否则若 $\mathbf{v}_1 \neq \mathbf{0}$，并且存在不全为零的权重 $c_1, \cdots, c_p$，使得

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$$

令 j 为满足 $c_j \neq \mathbf{0}$ 的最大下标. 若 $j = 1$，则 $c_1\mathbf{v}_1 = \mathbf{0}$，与 $\mathbf{v}_1 \neq \mathbf{0}$ 矛盾. 所以 $j > 1$，且

$$c_1\mathbf{v}_1 + \cdots + c_j\mathbf{v}_j + 0\mathbf{v}_{j+1} + \cdots + 0\mathbf{v}_p = \mathbf{0}$$

$$c_j\mathbf{v}_j = -c_1\mathbf{v}_1 - \cdots - c_{j-1}\mathbf{v}_{j-1}$$

$$\mathbf{v}_j = \left(-\frac{c_1}{c_j}\right)\mathbf{v}_1 + \cdots + \left(-\frac{c_{j-1}}{c_j}\right)\mathbf{v}_{j-1}$$ ■

基础练习

设 $\mathbf{u} = \begin{bmatrix}3\\2\\-4\end{bmatrix}$, $\mathbf{v} = \begin{bmatrix}-6\\1\\7\end{bmatrix}$, $\mathbf{w} = \begin{bmatrix}0\\-5\\2\end{bmatrix}$, $\mathbf{z} = \begin{bmatrix}3\\7\\-5\end{bmatrix}$.

1. 集合 $\{\mathbf{u},\mathbf{v}\}$，$\{\mathbf{u},\mathbf{w}\}$，$\{\mathbf{u},\mathbf{z}\}$，$\{\mathbf{v},\mathbf{w}\}$，$\{\mathbf{v},\mathbf{z}\}$ 和 $\{\mathbf{w},\mathbf{z}\}$ 是否线性无关？为什么？
2. 基础练习 1 的答案是否表明 $\{\mathbf{u},\mathbf{v},\mathbf{w},\mathbf{z}\}$ 线性无关？
3. 为了判断 $\{\mathbf{u},\mathbf{v},\mathbf{w},\mathbf{z}\}$ 是否线性相关，检查 $\mathbf{w}$ 是否是 $\mathbf{u}$，$\mathbf{v}$，$\mathbf{z}$ 的线性组合是一个明智的方法吗？
4. $\{\mathbf{u},\mathbf{v},\mathbf{w},\mathbf{z}\}$ 是否线性相关？

70

习题 1.7

在习题 1 ~ 4 中，判断向量是否线性无关. 给出证明.

1. $\begin{bmatrix}5\\0\\0\end{bmatrix}, \begin{bmatrix}7\\2\\-6\end{bmatrix}, \begin{bmatrix}9\\4\\-8\end{bmatrix}$ 2. $\begin{bmatrix}0\\0\\2\end{bmatrix}, \begin{bmatrix}0\\5\\-8\end{bmatrix}, \begin{bmatrix}-3\\4\\1\end{bmatrix}$ 3. $\begin{bmatrix}1\\-3\end{bmatrix}, \begin{bmatrix}-3\\9\end{bmatrix}$ 4. $\begin{bmatrix}-1\\4\end{bmatrix}, \begin{bmatrix}-2\\-8\end{bmatrix}$

在习题 5 ~ 8 中，判断矩阵中的列是否能构成一个线性无关集合. 给出证明.

5. $\begin{bmatrix}0 & -8 & 5\\3 & -7 & 4\\-1 & 5 & -4\\1 & -3 & 2\end{bmatrix}$ 6. $\begin{bmatrix}-4 & -3 & 0\\0 & -1 & 4\\1 & 0 & 3\\5 & 4 & 6\end{bmatrix}$ 7. $\begin{bmatrix}1 & 4 & -3 & 0\\-2 & -7 & 5 & 1\\-4 & -5 & 7 & 5\end{bmatrix}$ 8. $\begin{bmatrix}1 & -3 & 3 & -2\\-3 & 7 & -1 & 2\\0 & 1 & -4 & 3\end{bmatrix}$

在习题 9 和 10 中，(a) h 取何值时，$\boldsymbol{v}_3$ 在 Span$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$中；(b) h 取何值时，$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$线性相关？给出证明.

9. $\boldsymbol{v}_1=\begin{bmatrix}1\\-3\\2\end{bmatrix}, \boldsymbol{v}_2=\begin{bmatrix}-3\\9\\-6\end{bmatrix}, \boldsymbol{v}_3=\begin{bmatrix}5\\-7\\h\end{bmatrix}$

10. $\boldsymbol{v}_1=\begin{bmatrix}1\\-5\\-3\end{bmatrix}, \boldsymbol{v}_2=\begin{bmatrix}-2\\10\\6\end{bmatrix}, \boldsymbol{v}_3=\begin{bmatrix}2\\-9\\h\end{bmatrix}$

在习题 11 ~ 14 中，找到使得向量组线性相关的数 h. 给出证明.

11. $\begin{bmatrix}1\\-1\\4\end{bmatrix}, \begin{bmatrix}3\\-5\\7\end{bmatrix}, \begin{bmatrix}-1\\5\\h\end{bmatrix}$

12. $\begin{bmatrix}2\\-4\\1\end{bmatrix}, \begin{bmatrix}-6\\7\\-3\end{bmatrix}, \begin{bmatrix}8\\h\\4\end{bmatrix}$

13. $\begin{bmatrix}1\\5\\-3\end{bmatrix}, \begin{bmatrix}-2\\-9\\6\end{bmatrix}, \begin{bmatrix}3\\h\\-9\end{bmatrix}$

14. $\begin{bmatrix}1\\-1\\-3\end{bmatrix}, \begin{bmatrix}-5\\7\\8\end{bmatrix}, \begin{bmatrix}1\\1\\h\end{bmatrix}$

通过观察，判断习题 15 ~ 20 中的向量集是否线性无关. 给出证明.

15. $\begin{bmatrix}5\\1\end{bmatrix}, \begin{bmatrix}2\\8\end{bmatrix}, \begin{bmatrix}1\\3\end{bmatrix}, \begin{bmatrix}-1\\7\end{bmatrix}$ 16. $\begin{bmatrix}4\\-2\\6\end{bmatrix}, \begin{bmatrix}6\\-3\\9\end{bmatrix}$ 17. $\begin{bmatrix}3\\5\\-1\end{bmatrix}, \begin{bmatrix}0\\0\\0\end{bmatrix}, \begin{bmatrix}-6\\5\\4\end{bmatrix}$

18. $\begin{bmatrix}4\\4\end{bmatrix}, \begin{bmatrix}-1\\3\end{bmatrix}, \begin{bmatrix}2\\5\end{bmatrix}, \begin{bmatrix}8\\1\end{bmatrix}$ 19. $\begin{bmatrix}-8\\12\\-4\end{bmatrix}, \begin{bmatrix}2\\-3\\-1\end{bmatrix}$ 20. $\begin{bmatrix}1\\4\\-7\end{bmatrix}, \begin{bmatrix}-2\\5\\3\end{bmatrix}, \begin{bmatrix}0\\0\\0\end{bmatrix}$

在习题 21 和 22 中，判断每个命题的真假，仔细阅读课本后说明理由.

21. a. 如果方程 $A\boldsymbol{x}=0$ 有平凡解，则 A 中的列线性无关.
    b. 如果 S 是一个线性相关集，则 S 中的每个向量都是其他向量的线性组合.
    c. 每个 4×5 矩阵中的列都是线性相关的.
    d. 如果 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 线性无关，且$\{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\}$线性相关，则 $\boldsymbol{z}$ 在 Span$\{\boldsymbol{x},\boldsymbol{y}\}$中.
22. a. 两个向量线性相关当且仅当它们同在一条过原点的直线上.
    b. 如果一个集合包含的向量个数少于向量中元素的个数，则该集合线性无关.
    c. 如果 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 线性无关，且 $\boldsymbol{z}$ 在 Span$\{\boldsymbol{x},\boldsymbol{y}\}$中，则$\{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\}$线性相关.
    d. 如果 $\mathbf{R}^n$中的一个集合线性相关，则该集合包含的向量个数要多于向量中包含的元素个数.

在习题 23 ~ 26 中，给出矩阵可能的阶梯形式.

23. A 是一个列线性无关的 3×3 矩阵.
24. A 是一个列线性相关的 2×2 矩阵.
25. A 是一个 4×2 矩阵，$A=[\boldsymbol{a}_1\ \ \boldsymbol{a}_2]$，$\boldsymbol{a}_2$不是 $\boldsymbol{a}_1$的倍数.

26. A 是一个 4×3 矩阵，$A=[a_1\quad a_2\quad a_3]$，$\{a_1, a_2\}$线性无关，且 a_3不在 Span$\{a_1, a_2\}$中.
27. 如果一个 7×5 矩阵中的列线性无关，那么矩阵中必须含有多少个主元列？为什么？
28. 如果一个 5×7 矩阵中的列张成 $\mathbf{R}^5$，那么矩阵中必须含有多少个主元列？为什么？
29. 构造 3×2 矩阵 A 和 B，使得 $Ax=0$ 仅有一个平凡解，$Bx=0$ 有一个非平凡解.

71

30. a. 填补命题中的空格："如果 A 为 $m\times n$ 矩阵，则 A 中的列线性无关当且仅当 A 有________个主元列."
    b. 解释为什么(a)中的命题是正确的.

习题 31 和 32 不需要用行变换来求解.［提示：把 $Ax=0$ 写成向量方程.］

31. 已知 $A=\begin{bmatrix} 2 & 3 & 5 \\ -5 & 1 & -4 \\ -3 & -1 & -4 \\ 1 & 0 & 1 \end{bmatrix}$，观察到第 3 列是前两列的和. 求出 $Ax=0$ 的一个非平凡解.
32. 已知 $A=\begin{bmatrix} 4 & 1 & 6 \\ -7 & 5 & 3 \\ 9 & -3 & 3 \end{bmatrix}$，观察到第 1 列加上第 2 列的两倍等于第 3 列. 求出 $Ax=0$ 的一个非平凡解.

习题 33 ~ 38 中的命题要么正确(在所有情况下)要么错误(至少有一个错误例子). 如果错误，可构造一个例子说明这个命题为什么错误. 这样的例子称为该命题的*反例*. 如果命题正确，给出证明(一个特例并不能解释命题为什么正确. 你要做的比习题 21 和 22 多).

33. 如果 $v_1, \cdots, v_4$是 $\mathbf{R}^4$中向量，且 $v_3=2v_1+v_2$，则$\{v_1, v_2, v_3, v_4\}$线性相关.
34. 如果 $v_1, \cdots, v_4$是 $\mathbf{R}^4$中向量，且 $v_3=0$，则$\{v_1, v_2, v_3, v_4\}$线性相关.
35. 如果 v_1, v_2是 $\mathbf{R}^4$中向量，且 v_2不是 v_1的数量倍数，则$\{v_1, v_2\}$线性无关.
36. 如果 $v_1, \cdots, v_4$是 $\mathbf{R}^4$中向量，且 v_3不是 v_1, v_2, v_4的线性组合，则$\{v_1, v_2, v_3, v_4\}$线性无关.
37. 如果 $v_1, \cdots, v_4$是 $\mathbf{R}^4$中向量，且$\{v_1, v_2, v_3\}$线性相关，则$\{v_1, v_2, v_3, v_4\}$也线性相关.
38. 如果 $v_1, \cdots, v_4$是 $\mathbf{R}^4$中线性无关的向量，则$\{v_1, v_2, v_3\}$也是线性无关的.［提示：考虑 $x_1v_1+x_2v_2+x_3v_3+0\cdot v_4=0$.］
39. 如果 A 为 $m\times n$ 矩阵，并且对 $\mathbf{R}^m$中的任意 b，方程 $Ax=b$ 至多有一个解. 利用线性无关的定义来解释为什么 A 中的列必定是线性无关的.
40. 假设 A 为含有 n 个主元列的 $m\times n$ 矩阵，解释为什么对 $\mathbf{R}^m$中的任意 b，方程 $Ax=b$ 至多有一个解.［提示：解释 $Ax=b$ 为什么没有无限多个解.］

[M]在习题 41 和 42 中，尽可能使用 A 中的列来构造一个矩阵 B，使得方程 $Bx=0$ 仅有一个平凡解. 通过求解 $Bx=0$ 来验证你的工作.

41. $A=\begin{bmatrix} 8 & -3 & 0 & -7 & 2 \\ -9 & 4 & 5 & 11 & -7 \\ 6 & -2 & 2 & -4 & 4 \\ 5 & -1 & 7 & 0 & 10 \end{bmatrix}$

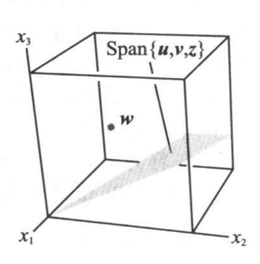
42. $A=\begin{bmatrix} 12 & 10 & -6 & -3 & 7 & 10 \\ -7 & -6 & 4 & 7 & -9 & 5 \\ 9 & 9 & -9 & -5 & 5 & -1 \\ -4 & -3 & 1 & 6 & -8 & 9 \\ 8 & 7 & -5 & -9 & 11 & -8 \end{bmatrix}$

43. [M]根据习题 41 中的 A 和 B，选出构造 B 时 A 中没有用到的列 v，判断 v 是否在 B 的列张成的空间中. 叙述计算过程.
44. [M]用习题 42 中的矩阵 A 和 B(B 是按指定要求构造的)重做习题 43，然后解释你的发现.

基础练习答案

1. 是，在每一种情形中，没有向量是其他向量的倍数. 因此每个集合都是线性无关的.
2. 不是. 基础练习1中的陈述没有表明任何关于$\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w},\boldsymbol{z}\}$线性无关的信息.
3. 不是. 在检验线性无关时，检查每个选出的向量是否是其他向量的线性组合，这个方法通常很粗劣. 因为，即使集合线性相关，选出的向量也不一定是其他向量的线性组合. 该练习中，$\boldsymbol{w}$并不是$\boldsymbol{u}$，$\boldsymbol{v}$，$\boldsymbol{z}$的线性组合.

72



4. 是. 根据定理8. 向量的个数(4)多于向量中元素的个数(3).

1.8　线性变换简介

矩阵方程$A\boldsymbol{x}=\boldsymbol{b}$及其对应的向量方程$x_1\boldsymbol{a}_1+\cdots+x_n\boldsymbol{a}_n=\boldsymbol{b}$只是记号不同，本质是相同的. 不过，线性代数的某些应用(例如计算机图形学和信号处理)中出现的矩阵方程$A\boldsymbol{x}=\boldsymbol{b}$与向量的线性组合并无直接联系，这时矩阵$A$被看成是“作用”在向量$\boldsymbol{x}$上的对象，初始向量$\boldsymbol{x}$乘以$A$后得到新向量$A\boldsymbol{x}$.

例如，方程

$$\underset{\uparrow\atop A}{\begin{bmatrix}4&-3&1&3\\2&0&5&1\end{bmatrix}}\underset{\uparrow\atop \boldsymbol{x}}{\begin{bmatrix}1\\1\\1\\1\end{bmatrix}}=\underset{\uparrow\atop \boldsymbol{b}}{\begin{bmatrix}5\\8\end{bmatrix}}\quad 和\quad \underset{\uparrow\atop A}{\begin{bmatrix}4&-3&1&3\\2&0&5&1\end{bmatrix}}\underset{\uparrow\atop \boldsymbol{u}}{\begin{bmatrix}1\\4\\-1\\3\end{bmatrix}}=\underset{\uparrow\atop \boldsymbol{0}}{\begin{bmatrix}0\\0\end{bmatrix}}$$

表示：$\boldsymbol{x}$乘以A得$\boldsymbol{b}$，$\boldsymbol{u}$乘以A得零向量. 见图1-32.

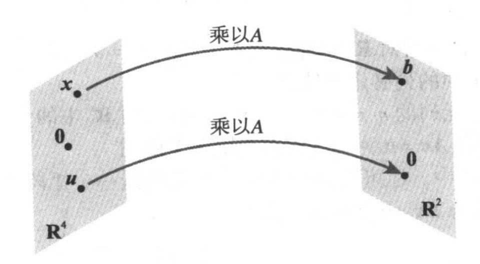


图1-32　通过矩阵乘法变换向量

根据这种新观点，求解方程$A\boldsymbol{x}=\boldsymbol{b}$相当于求$\mathbf{R}^4$中所有满足下列条件的向量$\boldsymbol{x}$：$\boldsymbol{x}$通过$A$的乘法变换能够得到$\mathbf{R}^2$中的$\boldsymbol{b}$.

将常见的函数概念(一个实数对应到另一实数)加以推广，我们可以把$\boldsymbol{x}$到$A\boldsymbol{x}$的对应看作是一个向量集到另一个向量集的函数.

一个$\mathbf{R}^n$到$\mathbf{R}^m$的**变换**(transformation)[**或函数**(function)、**映射**(mapping)]T是这样一个规则：为$\mathbf{R}^n$中每个向量$\boldsymbol{x}$指派$\mathbf{R}^m$中的一个向量$T(\boldsymbol{x})$. 集合$\mathbf{R}^n$称为变换T的

73

定义域(domain)，$\mathbf{R}^m$称为变换 T 的**上域**(codomain)．记号 T：$\mathbf{R}^n \to \mathbf{R}^m$表示定义域是 $\mathbf{R}^n$，上域是 $\mathbf{R}^m$．对 $\mathbf{R}^n$中的 $\boldsymbol{x}$，称 $\mathbf{R}^m$中的向量 $T(\boldsymbol{x})$为(变换 T 下)向量 $\boldsymbol{x}$ 的**像**(image)．全体像 $T(\boldsymbol{x})$ 的集合称为变换 T 的**值域**(range)．见图 1-33.

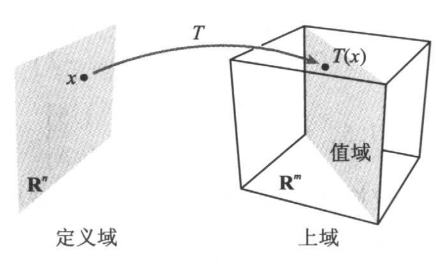


图 1-33 变换 T：$\mathbf{R}^n \to \mathbf{R}^m$的定义域、上域和值域

本节中的这个新概念很重要，因为用动态的眼光看待矩阵 - 向量乘法，对于理解线性代数的一些思想、构造随着时间变化的物理系统的数学模型是很关键的．这样的动态系统将在 1.10，4.8，4.9 以及整个第 5 章中讨论.

1.8.1 矩阵变换

本节的其余内容将讨论矩阵乘法所对应的映射．对任意 $\boldsymbol{x} \in \mathbf{R}^n$，算出 $A\boldsymbol{x}$ 作为 $T(\boldsymbol{x})$，其中 A 为 $m \times n$ 矩阵．为了简化，我们把这样的矩阵变换表示为：$\boldsymbol{x} \mapsto A\boldsymbol{x}$．$A$ 有 n 列时，变换 T 的定义域为 $\mathbf{R}^n$，A 的每一列有 m 个元素时，列变换 T 的上域为 $\mathbf{R}^m$．变换 T 的值域就是 A 中列的全体线性组合构成的集合.

【例题 1】 设 $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$，$\boldsymbol{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$，$\boldsymbol{b} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$，$\boldsymbol{c} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$，用 $T(\boldsymbol{x}) = A\boldsymbol{x}$ 定义一个变换 T：$\mathbf{R}^2 \to \mathbf{R}^3$，这样

$$T(\boldsymbol{x}) = A\boldsymbol{x} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{bmatrix}$$

a. 求出 $T(\boldsymbol{u})$，即 $\boldsymbol{u}$ 在变换 T 下的像.
b. 求 $\boldsymbol{x} \in \mathbf{R}^2$，$\boldsymbol{x}$ 在变换 T 下的像为 $\boldsymbol{b}$.
c. 是否存在多个 $\boldsymbol{x}$，其在变换 T 下的像均为 $\boldsymbol{b}$.
d. 判断 $\boldsymbol{c}$ 是否属于变换 T 的值域.

74

解：

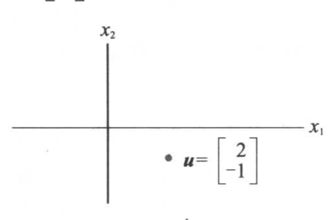
a. 计算（见图 1-34）

$$T(\boldsymbol{u}) = A\boldsymbol{u} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}$$

b. 求解 $T(\boldsymbol{x}) = \boldsymbol{b}$ 中的 $\boldsymbol{x}$，即求解 $A\boldsymbol{x} = \boldsymbol{b}$，或者

$$\begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix} \quad (1)$$

用 1.4 节中的方法，对增广矩阵进行行化简：



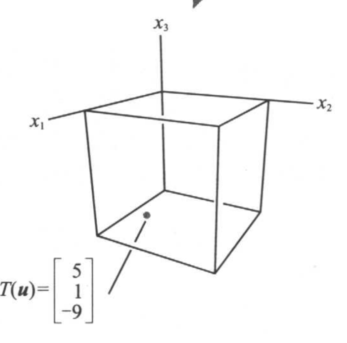


图 1-34

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & -0.5 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1.5 \\ 0 & 1 & -0.5 \\ 0 & 0 & 0 \end{bmatrix} \quad (2)$$

因此，$x_1 = 1.5$，$x_2 = -0.5$，$\boldsymbol{x} = \begin{bmatrix} 1.5 \\ -0.5 \end{bmatrix}$. $\boldsymbol{x}$ 在变换 T 下的像就是给定的向量 $\boldsymbol{b}$.

c. 在变换 T 下像为 $\boldsymbol{b}$ 的向量 $\boldsymbol{x}$ 必须满足(1). 根据(2)，方程(1)显然有唯一解. 所以这样的 $\boldsymbol{x}$ 只有一个.

d. 如果 $\boldsymbol{c}$ 为 $\mathbf{R}^2$ 中某个 $\boldsymbol{x}$ 的像，即存在 $\boldsymbol{x}$ 使得 $\boldsymbol{c} = T(\boldsymbol{x})$，则向量 $\boldsymbol{c}$ 属于 T 的值域. 因此本题相当于判断方程组 $A\boldsymbol{x} = \boldsymbol{c}$ 是否相容. 为了找到这个答案，对增广矩阵进行行化简：

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 14 & -7 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & -35 \end{bmatrix}$$

第三个方程为 $0 = -35$，说明该方程组不相容. 因此 $\boldsymbol{c}$ 不属于变换 T 的值域. ■

例题1(c)中的问题就是线性方程组的*唯一性*问题，只不过这里使用矩阵变换加以陈述：$\boldsymbol{b}$ 是否是 $\mathbf{R}^n$ 中唯一一个 $\boldsymbol{x}$ 的像？类似地，例题1(d)讨论的是*存在性*问题：是否存在像为 $\boldsymbol{c}$ 的 $\boldsymbol{x}$？

接下来的两个矩阵变换可以看成几何变换. 这时，矩阵的动态变化理解为一个向量变换成另一个向量. 2.7节还将给出与计算机图形学有关的若干有趣的例子.

75

【例题2】　如果 $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$，则变换 $\boldsymbol{x} \mapsto A\boldsymbol{x}$ 把 $\mathbf{R}^3$ 中的点投影到 x_1x_2 平面，因为

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$$

见图1-35.

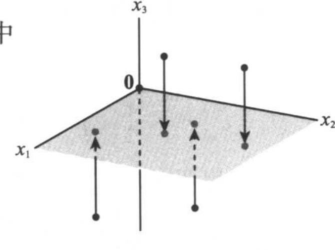


图1-35　投影变换

■

【例题3】　设 $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$. 由 $T(\boldsymbol{x}) = A\boldsymbol{x}$ 定义的变换 $T: \mathbf{R}^2 \to \mathbf{R}^2$ 称作**剪切变换**(shear transformation). 可以证明，如果将 T 作用于图1-36所示的 2×2 正方形中的每一点，则像的集合构成平行四边形. 证明的关键是说明变换 T 把直线段映到直线段(见习题27)，然后验证正方形的顶点映到平行四边形的顶点上. 例如，点 $\boldsymbol{u} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ 的像为 $T(\boldsymbol{u}) = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$，$\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ 的像为 $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$. 变换 T 使正方形发生形变，

使其底部固定不动，而顶部被推向右边. 剪切变换常见于物理学、几何学和晶体学当中.

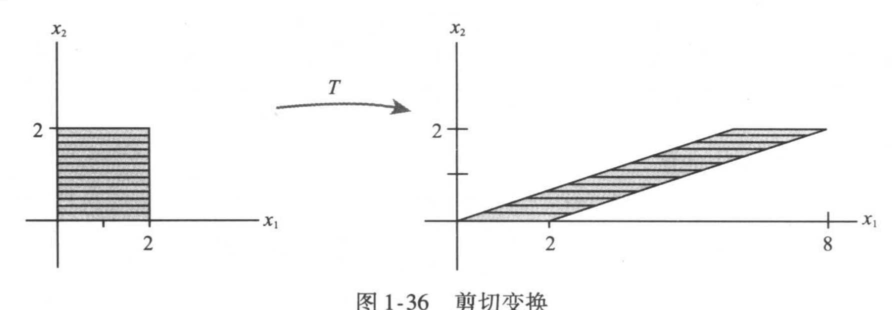


图 1-36 剪切变换

■

1.8.2 线性变换

1.4 节的定理 5 表明，如果 $\boldsymbol{A}$ 为 $m\times n$ 矩阵，对任意 $\boldsymbol{u}$，$\boldsymbol{v}\in\mathbf{R}^n$ 以及任意数量 c，变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 满足性质

$$A(\boldsymbol{u}+\boldsymbol{v})=A\boldsymbol{u}+A\boldsymbol{v} \quad 和 \quad A(c\boldsymbol{u})=cA\boldsymbol{u}$$

这些性质确定了线性代数中最重要的一类变换.

76

【定义】 一个变换(或映射) T **是线性的**(linear)，如果：

(i)对 T 定义域中任意的 $\boldsymbol{u}$，$\boldsymbol{v}$，都有 $T(\boldsymbol{u}+\boldsymbol{v})=T(\boldsymbol{u})+T(\boldsymbol{v})$；

(ii)对任意的 $\boldsymbol{u}$ 和数量 c，都有 $T(c\boldsymbol{u})=cT(\boldsymbol{u})$.

矩阵变换都是线性变换其他一些重要的线性变换将在第 4 章和第 5 章中讨论.

线性变换保持向量的加法和数乘运算. 性质(i)说的是：在 $\mathbf{R}^n$ 中先算 $\boldsymbol{u}$ 加 $\boldsymbol{v}$，然后执行变换 T，得到的结果 $T(\boldsymbol{u}+\boldsymbol{v})$ 与在 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 上先分别执行变换 T，再把 $\mathbf{R}^m$ 中的 $T(\boldsymbol{u})$ 和 $T(\boldsymbol{v})$ 相加所得到的结果一样. 由这两个性质很容易得到下述有用的事实.

如果 T 是一个线性变换，则对 T 定义域中任意向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 以及任意数量 c 和 d，都有：

$$T(\mathbf{0})=\mathbf{0} \tag{3}$$

且

$$T(c\boldsymbol{u}+d\boldsymbol{v})=cT(\boldsymbol{u})+dT(\boldsymbol{v}) \tag{4}$$

性质(3)由定义中的(ii)得到，因为 $T(\mathbf{0})=T(0\boldsymbol{u})=0T(\boldsymbol{u})=\mathbf{0}$. 性质(4)需要用到(i)和(ii)：

$$T(c\boldsymbol{u}+d\boldsymbol{v})=T(c\boldsymbol{u})+T(d\boldsymbol{v})=cT(\boldsymbol{u})+dT(\boldsymbol{v})$$

注意，如果一个变换对于任意的 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 以及 c 和 d 都满足(4)，这个变换一定是线性的(令 $c=d=1$ 可验证它保持加法；令 $d=0$ 可验证它保持数乘). 反复应用(4)可得到一个有用的推广：

$$T(c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p)=c_1T(\boldsymbol{v}_1)+\cdots+c_pT(\boldsymbol{v}_p) \tag{5}$$

在工程学和物理学中，(5)被称作叠加原理. 在信号系统中，令输入的单个信号

为 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$，对应的响应信号为 $T(\boldsymbol{v}_1),\cdots,T(\boldsymbol{v}_p)$. 如果将信号系统的输入表示成单个信号的线性组合，则系统的响应就是单个信号响应的同一线性组合，这时称该信号系统满足叠加原理. 我们将在第4章中继续讨论这一思想.

【例题4】 已知一个数量 r，定义变换 T：$\mathbf{R}^2\to\mathbf{R}^2$ 为 $T(\boldsymbol{x})=r\boldsymbol{x}$. $0\leq r\leq 1$ 时，称 T 为**压缩变换**(contraction)；$r>1$ 时，称 T 为**膨胀变换**(dilation). 设 $r=3$，证明 T 是线性变换.

77

解：设 $\boldsymbol{u}$，$\boldsymbol{v}$ 是 $\mathbf{R}^2$ 中向量，c，d 为数量. 则

$$\begin{aligned} T(c\boldsymbol{u}+d\boldsymbol{v}) &= 3(c\boldsymbol{u}+d\boldsymbol{v}) && T\text{的定义} \\ &= 3c\boldsymbol{u}+3d\boldsymbol{v} && \\ &= c(3\boldsymbol{u})+d(3\boldsymbol{v}) && \text{向量算术} \\ &= cT(\boldsymbol{u})+dT(\boldsymbol{v}) \end{aligned}$$

因此 T 满足(4)，是一线性变换. 见图1-37.

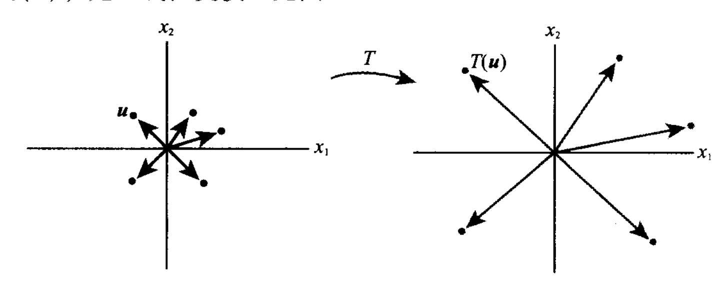


图1-37　一个膨胀变换

■

【例题5】 定义线性变换 T：$\mathbf{R}^2\to\mathbf{R}^2$ 为

$$T(\boldsymbol{x})=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}-x_2\\ x_1\end{bmatrix}$$

求 $\boldsymbol{u}=\begin{bmatrix}4\\1\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}2\\3\end{bmatrix}$ 和 $\boldsymbol{u}+\boldsymbol{v}=\begin{bmatrix}6\\4\end{bmatrix}$ 在变换 T 下的像.

解：

$$T(\boldsymbol{u})=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}4\\ 1\end{bmatrix}=\begin{bmatrix}-1\\ 4\end{bmatrix},\quad T(\boldsymbol{v})=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}2\\ 3\end{bmatrix}=\begin{bmatrix}-3\\ 2\end{bmatrix},$$

$$T(\boldsymbol{u}+\boldsymbol{v})=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}6\\ 4\end{bmatrix}=\begin{bmatrix}-4\\ 6\end{bmatrix}$$

注意，$T(\boldsymbol{u}+\boldsymbol{v})$ 显然等于 $T(\boldsymbol{u})+T(\boldsymbol{v})$. 图1-38表明变换 T 将 $\boldsymbol{u}$，$\boldsymbol{v}$ 和 $\boldsymbol{u}+\boldsymbol{v}$ 绕原点逆时针旋转了90度. 事实上，变换 T 把 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 确定的平行四边形变换成 $T(\boldsymbol{u})$ 和 $T(\boldsymbol{v})$ 确定的平行四边形.（见习题28.）

78

最后这道例题与几何无关；它演示了线性映射怎样把一种类型的数据变换成另一种类型.

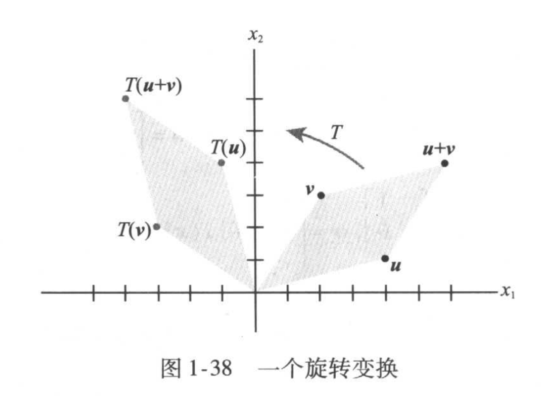


图 1-38 一个旋转变换

■

【例题 6】 一个公司生产两种产品 B 和 C. 使用 1.3 节例题 7 中的数据，我们构造"单位成本"矩阵 $U=[\boldsymbol{b}\quad \boldsymbol{c}]$，其列向量表示产品的"一美元产出成本".

$$\begin{array}{c} \text{产品} \\ \begin{array}{cc} \text{B} & \text{C} \end{array} \end{array}$$

$$U=\begin{bmatrix}0.45 & 0.40\\ 0.25 & 0.35\\ 0.15 & 0.15\end{bmatrix}\begin{array}{l}\text{原材料}\\ \text{劳动力}\\ \text{管理费}\end{array}$$

设 $\boldsymbol{x}=(x_1,x_2)$ 是"产量"向量，代表 $\boldsymbol{x}_1$ 美元的产品 B 和 $\boldsymbol{x}_2$ 美元的产品 C，此外定义变换 T：$\mathbf{R}^2\to\mathbf{R}^3$ 如下：

$$T(\boldsymbol{x})=U\boldsymbol{x}=x_1\begin{bmatrix}0.45\\ 0.25\\ 0.15\end{bmatrix}+x_2\begin{bmatrix}0.40\\ 0.35\\ 0.15\end{bmatrix}=\begin{bmatrix}\text{原材料总成本}\\ \text{劳动力总成本}\\ \text{管理费总成本}\end{bmatrix}$$

映射 T 把一列产量(单位:美元)转换成一列总成本. 这个映射的线性性体现在以下两个方面：

1. 如果产量增加到 4 倍，即从 $\boldsymbol{x}$ 增加到 $4\boldsymbol{x}$，则成本将增长同样的倍数，从 $T(\boldsymbol{x})$ 到 $4T(\boldsymbol{x})$.

2. 如果 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 为产量向量，则与产量之和 $\boldsymbol{x}+\boldsymbol{y}$ 对应的总成本向量恰好是成本向量 $T(\boldsymbol{x})$ 与 $T(\boldsymbol{y})$ 的和. ■

基础练习

1. 假设 T：$\mathbf{R}^5\to\mathbf{R}^2$，且对任意 $\boldsymbol{x}\in\mathbf{R}^5$ 都有 $T(\boldsymbol{x})=A\boldsymbol{x}$，其中 A 为矩阵. 那么 A 中有几行几列？

2. 设 $A=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$. 给出变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的几何表示.

3. 从 $\mathbf{0}$ 到向量 $\boldsymbol{u}$ 的线段是形如 $t\boldsymbol{u}$，$0\leqslant t\leqslant 1$ 的点所构成的集合. 证明线性变换 T 把该线段映为从 $\mathbf{0}$ 到 $T(\boldsymbol{u})$ 的线段.

习题 1.8

1. 设$A=\begin{bmatrix}2&0\\0&2\end{bmatrix}$，用$T(\boldsymbol{x})=A\boldsymbol{x}$定义$T$：$\mathbf{R}^2\to\mathbf{R}^2$. 求出$\boldsymbol{u}=\begin{bmatrix}1\\-3\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}a\\b\end{bmatrix}$在变换下$T$的像.

2. 设$A=\begin{bmatrix}0.5&0&0\\0&0.5&0\\0&0&0.5\end{bmatrix}$，$\boldsymbol{u}=\begin{bmatrix}1\\0\\-4\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}a\\b\\c\end{bmatrix}$，用$T(\boldsymbol{x})=A\boldsymbol{x}$定义$T$：$\mathbf{R}^3\to\mathbf{R}^3$，求出$T(\boldsymbol{u})$，

79

$T(\boldsymbol{v})$.

在习题3～6中，T定义为$T(\boldsymbol{x})=A\boldsymbol{x}$，求出向量$\boldsymbol{x}$，它在$T$下的像为$\boldsymbol{b}$，判断$\boldsymbol{x}$是否唯一.

3. $A=\begin{bmatrix}1&0&-2\\-2&1&6\\3&-2&-5\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}-1\\7\\-3\end{bmatrix}$

4. $A=\begin{bmatrix}1&-3&2\\0&1&-4\\3&-5&-9\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}6\\-7\\-9\end{bmatrix}$

5. $A=\begin{bmatrix}1&-5&-7\\-3&7&5\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}-2\\-2\end{bmatrix}$

6. $A=\begin{bmatrix}1&-2&1\\3&-4&5\\0&1&1\\-3&5&-4\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}1\\9\\3\\-6\end{bmatrix}$

7. 设A为一个6×5矩阵，如果$T(\boldsymbol{x})=A\boldsymbol{x}$定义$T$：$\mathbf{R}^a\to\mathbf{R}^b$，$a$和$b$应该是多少？

8. 如果$T(\boldsymbol{x})=A\boldsymbol{x}$定义了一个从$\mathbf{R}^4$到$\mathbf{R}^5$的映射，矩阵$A$应该有几行几列？

对习题9和10，根据给定的矩阵A，找出$\mathbf{R}^4$中通过变换$\boldsymbol{x}\mapsto A\boldsymbol{x}$映射到零向量的所有$\boldsymbol{x}$.

9. $A=\begin{bmatrix}1&-4&7&-5\\0&1&-4&3\\2&-6&6&-4\end{bmatrix}$

10. $A=\begin{bmatrix}1&3&9&2\\1&0&3&-4\\0&1&2&3\\-2&3&0&5\end{bmatrix}$

11. 设$\boldsymbol{b}=\begin{bmatrix}-1\\1\\0\end{bmatrix}$，且$A$为习题9中的矩阵，$\boldsymbol{b}$是否在线性变换$\boldsymbol{x}\mapsto A\boldsymbol{x}$的值域中？为什么？

12. 设$\boldsymbol{b}=\begin{bmatrix}-1\\3\\-1\\4\end{bmatrix}$，且$A$为习题10中的矩阵，$\boldsymbol{b}$是否在线性变换$\boldsymbol{x}\mapsto A\boldsymbol{x}$的值域中？为什么？

在习题13～16中，$\boldsymbol{u}=\begin{bmatrix}5\\2\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}-2\\4\end{bmatrix}$，根据给定的变换$T$，用几何方法表示出$T$怎样作用于$\mathbf{R}^2$中每一点$\boldsymbol{x}$(在每个习题后作图).

13. $T(\boldsymbol{x})=\begin{bmatrix}-1&0\\0&-1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$

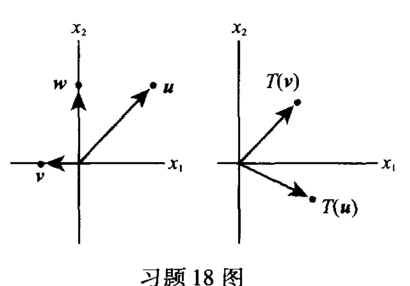
14. $T(\boldsymbol{x})=\begin{bmatrix}0.5&0\\0&0.5\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$

15. $T(\boldsymbol{x})=\begin{bmatrix}0&0\\0&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$

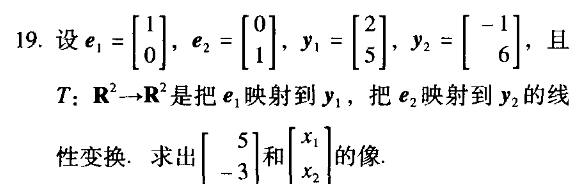
16. $T(\boldsymbol{x})=\begin{bmatrix}0&1\\1&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$

17. 设 T: $\mathbf{R}^2 \to \mathbf{R}^2$ 是线性变换，它把 $\boldsymbol{u} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$ 映射到 $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$，把 $\boldsymbol{v} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ 映射到 $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$. 利用变换 T 的线性性求出 $3\boldsymbol{u}$，$2\boldsymbol{v}$，$3\boldsymbol{u}+2\boldsymbol{v}$ 在变换 T 下的像.

18. 右图中标出向量 $\boldsymbol{u}$，$\boldsymbol{v}$ 和 $\boldsymbol{w}$，以及在变换 T: $\mathbf{R}^2 \to \mathbf{R}^2$ 作用下的 $T(\boldsymbol{u})$ 和 $T(\boldsymbol{v})$. 认真参照该图，尽可能准确地画出 $T(\boldsymbol{w})$. [提示：首先把 $\boldsymbol{w}$ 写成 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 线性组合的形式.]



习题 18 图


19. 设 $\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$，$\boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$，$\boldsymbol{y}_1 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$，$\boldsymbol{y}_2 = \begin{bmatrix} -1 \\ 6 \end{bmatrix}$，且 T: $\mathbf{R}^2 \to \mathbf{R}^2$ 是把 $\boldsymbol{e}_1$ 映射到 $\boldsymbol{y}_1$，把 $\boldsymbol{e}_2$ 映射到 $\boldsymbol{y}_2$ 的线性变换. 求出 $\begin{bmatrix} 5 \\ -3 \end{bmatrix}$ 和 $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 的像.

20. 设 $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$，$\boldsymbol{v}_1 = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$，$\boldsymbol{v}_2 = \begin{bmatrix} 7 \\ -3 \end{bmatrix}$，且 T: $\mathbf{R}^2 \to \mathbf{R}^2$ 为把 $\boldsymbol{x}$ 映射到 $x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2$ 的线性变换. 求出矩阵 $\boldsymbol{A}$，使得对于任意的 $\boldsymbol{x}$，$T(\boldsymbol{x})$ 都等于 $\boldsymbol{A}\boldsymbol{x}$.

在习题 21 和 22 中，判断各命题的真假，并说明理由.

21. a. 线性变换是函数的一个特殊形式.
    b. 如果 $\boldsymbol{A}$ 是一个 3×5 矩阵，且 $\boldsymbol{T}$ 是用 $\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ 定义的一个变换，则 T 的值域为 $\mathbf{R}^3$.
    c. 如果 $\boldsymbol{A}$ 为 $m \times n$ 矩阵，则变换 $\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$ 的值域为 $\mathbf{R}^m$.
    d. 每个线性变换都是矩阵变换.
    e. 变换 T 是线性的当且仅当对于 T 的定义域中的任意 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$，以及任意数量 c_1 和 c_2，都有 $T(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) = c_1T(\boldsymbol{v}_1) + c_2T(\boldsymbol{v}_2)$.

80

22. a. 每个矩阵变换都是线性变换.
    b. 变换 $\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$ 的上域是 $\boldsymbol{A}$ 中列的所有线性组合构成的集合.
    c. 如果 T: $\mathbf{R}^n \to \mathbf{R}^m$ 是一个线性变换，且 $\boldsymbol{c}$ 在 $\mathbf{R}^m$ 中，则唯一性问题是"$\boldsymbol{c}$ 是否在 T 的定义域中?"
    d. 线性变换保持向量加法和数乘运算.
    e. 叠加原理是线性变换在物理学中的一种表述.

23. 设线性变换 T: $\mathbf{R}^2 \to \mathbf{R}^2$ 将每一点做关于 x_1 轴的反射(见基础练习 2). 作与图 1-38 类似的两个图，表示出线性变换的性质(i)和(ii).

24. 假设向量 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p$ 张成 $\mathbf{R}^n$，设 T: $\mathbf{R}^n \to \mathbf{R}^n$ 为一个线性变换. 如果对 $i = 1, \cdots, p$，有 $T(\boldsymbol{v}_i) = \mathbf{0}$. 证明 T 为零变换. 即，若 $\boldsymbol{x}$ 是 $\mathbf{R}^n$ 中向量，则 $T(\boldsymbol{x}) = \mathbf{0}$.

25. 已知 $\mathbf{R}^n$ 中的向量 $\boldsymbol{p}$，且 $\boldsymbol{v} \neq 0$，沿 $\boldsymbol{v}$ 的方向过 $\boldsymbol{p}$ 的直线有参数方程形式 $\boldsymbol{x} = \boldsymbol{p} + t\boldsymbol{v}$. 证明线性变换 T: $\mathbf{R}^n \to \mathbf{R}^n$ 把这条直线映射到另一条直线或单点(退化的直线)上.

26. 设 $\mathbf{R}^3$ 中的向量 $\boldsymbol{u}$，$\boldsymbol{v}$ 线性无关，且 P 为过点 $\boldsymbol{u}$，$\boldsymbol{v}$，0 的平面. P 的参数方程为 $\boldsymbol{x} = s\boldsymbol{u} + t\boldsymbol{v}$($s$，$t$ 在 $\mathbf{R}$ 中). 证明线性变换 T: $\mathbf{R}^3 \to \mathbf{R}^3$ 把平面 P 影射到一个过 0 点的平面上，或者一条过 0 点的直线上，或者仅仅是 $\mathbf{R}^3$ 中的原点. 为了使平面 P 的像还是一个平面，$T(\boldsymbol{u})$，$T(\boldsymbol{v})$ 应该具备怎样的条件?

27. a. 证明过 $\mathbf{R}^n$ 中向量 $\boldsymbol{p}$，$\boldsymbol{q}$ 的直线可以写成参数形式 $\boldsymbol{x} = (1-t)\boldsymbol{p} + t\boldsymbol{q}$. (参照 1.5 节习题 21 和 22.)
    b. 从 $\boldsymbol{p}$ 到 $\boldsymbol{q}$ 的线段方程形如 $(1-t)\boldsymbol{p} + t\boldsymbol{q}$，$0 \leqslant t \leqslant 1$ 的点的集合(如

$(t=1)$ $\boldsymbol{q}$　$(1-t)\boldsymbol{p}+t\boldsymbol{q}$　$\boldsymbol{x}$　$(t=0)$ $\boldsymbol{p}$

习题 27 图

习题27图所示). 证明线性变换 T 把该线段映射到另一条线段或者单点上.

28. 设 $\boldsymbol{u}$, $\boldsymbol{v}$ 为 $\mathbf{R}^n$ 中向量. 可以证明：由 $\boldsymbol{u}$, $\boldsymbol{v}$ 确定的平行四边形 P 上所有的点都有 $a\boldsymbol{u}+b\boldsymbol{v}$ 的形式，这里 $0\leqslant a\leqslant 1$, $0\leqslant b\leqslant 1$. 设 T: $\mathbf{R}^n\to\mathbf{R}^m$ 是一个线性变换. 解释在变换 T 下，为什么 P 中点的像在 $T(\boldsymbol{u})$, $T(\boldsymbol{v})$ 确定的平行四边形中.
29. 用 $f(x)=mx+b$ 定义变换 f: $\mathbf{R}\to\mathbf{R}$.
    a. 证明 $b=0$ 时，f 为线性变换.
    b. 找出线性变换的一条性质，它在 $b\neq 0$ 时不成立.
    c. 为什么 f 被称为线性函数?
30. 仿射变换 T: $\mathbf{R}^n\to\mathbf{R}^m$ 有这样的形式：$T(\boldsymbol{x})=A\boldsymbol{x}+\boldsymbol{b}$, A 为 $m\times n$ 矩阵，且 $\boldsymbol{b}$ 在 $\mathbf{R}^m$ 中. 证明当 $\boldsymbol{b}\neq 0$ 时，T 不是线性变换. 仿射变换在计算机绘图中很重要.
31. 设 T: $\mathbf{R}^n\to\mathbf{R}^m$ 为一个线性变换，且 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 为 $\mathbf{R}^n$ 中线性相关集合. 解释为什么集合 $\{T(\boldsymbol{v}_1),T(\boldsymbol{v}_2),T(\boldsymbol{v}_3)\}$ 线性相关.

在习题32～36中，列向量写成了行的形式，如 $\boldsymbol{x}=(x_1,x_2)$, $T(\boldsymbol{x})$ 写成 $T(x_1,x_2)$.

32. 证明用 $T(x_1,x_2)=(4x_1-2x_2,3|x_2|)$ 定义的变换 T 不是线性变换.
33. 证明用 $T(x_1,x_2)=(2x_1-3x_2,x_1+4,5x_2)$ 定义的变换 T 不是线性变换.
34. 设 T: $\mathbf{R}^n\to\mathbf{R}^m$ 是一个线性变换. 证明如果变换 T 把两个线性无关的向量映射到一个线性相关集上，则方程 $T(\boldsymbol{x})=0$ 有非平凡解. [提示：设 $\mathbf{R}^n$ 中线性无关的向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$, $T(\boldsymbol{u})$ 和 $T(\boldsymbol{v})$ 线性相关. 则存在不全为零的数量 c_1、c_2，使得 $c_1T(\boldsymbol{u})+c_2T(\boldsymbol{v})=0$. 利用这个等式.]
35. 设变换 T: $\mathbf{R}^3\to\mathbf{R}^3$ 将每个向量 $\boldsymbol{x}=(x_1,x_2,x_3)$ 做关于平面 $x_3=0$ 的反射，得到 $T(\boldsymbol{x})=(x_1,x_2,-x_3)$. 证明 T 是一个线性变换. [用例题4中的思想.]
36. 设变换 T: $\mathbf{R}^3\to\mathbf{R}^3$ 把每个向量 $\boldsymbol{x}=(x_1,x_2,x_3)$ 投影到平面 $x_2=0$ 上，得到 $T(\boldsymbol{x})=(x_1,0,x_3)$. 证明 T 是一个线性变换.

[M]在习题37和38中，已知矩阵确定了一个线性变换 T. 求出所有使得 $T(\boldsymbol{x})=\boldsymbol{0}$ 的 $\boldsymbol{x}$.

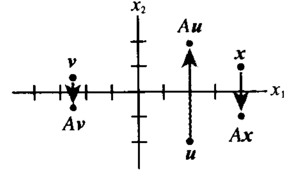
81

37. $\begin{bmatrix} 4 & -2 & 5 & -5 \\ -9 & 7 & -8 & 0 \\ -6 & 4 & 5 & 3 \\ 5 & -3 & 8 & -4 \end{bmatrix}$

38. $\begin{bmatrix} -9 & -4 & -9 & 4 \\ 5 & -8 & -7 & 6 \\ 7 & 11 & 16 & -9 \\ 9 & -7 & -4 & 5 \end{bmatrix}$

39. [M]设 $\boldsymbol{b}=\begin{bmatrix} 7 \\ 5 \\ 9 \\ 7 \end{bmatrix}$，且 A 为习题37中的矩阵. $\boldsymbol{b}$ 是否在变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的值域中? 如果在，求出在该变换下像为 $\boldsymbol{b}$ 的 $\boldsymbol{x}$.

40. [M]$\boldsymbol{b}=\begin{bmatrix} -7 \\ -7 \\ 13 \\ -5 \end{bmatrix}$，且 A 为习题38中的矩阵. $\boldsymbol{b}$ 是否在变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的值域中? 如果在，求出在该变换下像为 $\boldsymbol{b}$ 的 $\boldsymbol{x}$.



练习2图 变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$

基础练习答案

1. 为使 $A\boldsymbol{x}$ 可定义，A 必须有5列. 为使 T 的上域为 $\mathbf{R}^2$，A 必须有2行.
2. 在图纸上绘制一些随机点(向量)，看看将出现什么情况. 例如点$(4,1)$映射到$(4,-1)$. 变

换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 将点做关于 x 轴(或 x_1 轴)的反射.

3. 设 $\boldsymbol{x}=t\boldsymbol{u}$，$0\leqslant t\leqslant 1$. 由于 T 是线性的，则 $T(t\boldsymbol{u})=tT(\boldsymbol{u})$ 是 $\boldsymbol{0}$ 和 $T(\boldsymbol{u})$ 之间线段上的一点.

1.9 线性变换的矩阵

一个线性变换 T，无论是用图示或文字描述，我们都希望得到 $T(\boldsymbol{x})$ 的“计算式”. 下面的讨论表明，从 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 的每个线性变换事实上都是一个矩阵变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$，并且 T 的主要性质与我们所熟知的 A 的性质密切相关. 求 A 的关键，是注意 T 完全由它在 $n\times n$ 单位矩阵 I_n 列上的作用所确定.

x_2

$e_2=\begin{bmatrix}0\\1\end{bmatrix}$

$e_1=\begin{bmatrix}1\\0\end{bmatrix}$

x_1

图 1-39

【例题 1】 $I_2=\begin{bmatrix}1&0\\0&1\end{bmatrix}$，$I_2$ 中的列为 $\boldsymbol{e}_1=\begin{bmatrix}1\\0\end{bmatrix}$，$\boldsymbol{e}_2=\begin{bmatrix}0\\1\end{bmatrix}$ (见图 1-39). 假设 T 是从 $\mathbf{R}^2$ 到 $\mathbf{R}^3$ 的线性变换：

$$T(\boldsymbol{e}_1)=\begin{bmatrix}5\\-7\\2\end{bmatrix},\quad T(\boldsymbol{e}_2)=\begin{bmatrix}-3\\8\\0\end{bmatrix}$$

除此之外没有其他信息，求任意 $\boldsymbol{x}\in\mathbf{R}^2$ 的像的公式.

解：

$$\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\end{bmatrix}=x_1\begin{bmatrix}1\\0\end{bmatrix}+x_2\begin{bmatrix}0\\1\end{bmatrix}=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2 \tag{1}$$

82

由于 T 是线性变换，

$$T(\boldsymbol{x})=x_1T(\boldsymbol{e}_1)+x_2T(\boldsymbol{e}_2)=x_1\begin{bmatrix}5\\-7\\2\end{bmatrix}+x_2\begin{bmatrix}-3\\8\\0\end{bmatrix}=\begin{bmatrix}5x_1-3x_2\\-7x_1+8x_2\\2x_1+0\end{bmatrix} \tag{2}$$

■

(1)到(2)的这一步解释了为什么已知 $T(\boldsymbol{e}_1)$，$T(\boldsymbol{e}_2)$ 就足以对任意 $\boldsymbol{x}$ 确定 $T(\boldsymbol{x})$. 此外，由于(2)把线性变换 T 表示成向量的线性组合形式，可以把这些向量作为 A 中的列，把(2)写成

$$T(\boldsymbol{x})=[T(\boldsymbol{e}_1)\quad T(\boldsymbol{e}_2)]\begin{bmatrix}x_1\\x_2\end{bmatrix}=A\boldsymbol{x}$$

【定理 10】 设 T：$\mathbf{R}^n\to\mathbf{R}^m$ 为线性变换. 则存在唯一一个矩阵 A，使得

$$\text{对任意 } \boldsymbol{x}\in\mathbf{R}^n\text{，有 } T(\boldsymbol{x})=A\boldsymbol{x}$$

事实上，A 是第 j 列为向量 $T(\boldsymbol{e}_j)$ 的 $m\times n$ 矩阵，其中 $\boldsymbol{e}_j$ 为 $\mathbf{R}^n$ 中单位矩阵的第 j 列，即：

$$A=[T(\boldsymbol{e}_1)\cdots T(\boldsymbol{e}_n)] \tag{3}$$

证明： $\boldsymbol{x}=\boldsymbol{I}_n\boldsymbol{x}=[\boldsymbol{e}_1\cdots\boldsymbol{e}_n]\boldsymbol{x}=x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_n$，利用 T 的线性性

$$T(\boldsymbol{x})=T(x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_n)=x_1T(\boldsymbol{e}_1)+\cdots+x_nT(\boldsymbol{e}_n)=[T(\boldsymbol{e}_1)\cdots T(\boldsymbol{e}_n)]\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}=\boldsymbol{A}\boldsymbol{x}$$

$\boldsymbol{A}$ 的唯一性将在习题 33 中证明. ■

(3)中的矩阵 $\boldsymbol{A}$ 称为**线性变换 T 的标准矩阵**.

我们现在知道从 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 的每个线性变换都是一个矩阵变换，反之亦然. 线性变换这一概念强调它作为映射的性质，而矩阵变换则描述了映射是怎样实现的，如下面例题所示.

83

【例题 2】 对 $\boldsymbol{x}\in\mathbf{R}^2$，求膨胀变换 $T(\boldsymbol{x})=3\boldsymbol{x}$ 的标准矩阵.

解：

$$T(e_1)=3e_1=\begin{bmatrix}3\\0\end{bmatrix}\quad 且\quad T(e_2)=3e_2=\begin{bmatrix}0\\3\end{bmatrix}$$

$$A=\begin{bmatrix}3&0\\0&3\end{bmatrix}$$

■

【例题 3】 设 T：$\mathbf{R}^2\to\mathbf{R}^2$ 为 $\mathbf{R}^2$ 中将每一点绕原点逆时针旋转 φ 角度的变换. 我们可以用几何方法证明该变换是线性的(见 1.8 节图 1-38). 求该变换的标准矩阵 $\boldsymbol{A}$.

解： $\begin{bmatrix}1\\0\end{bmatrix}$旋转到$\begin{bmatrix}\cos\varphi\\ \sin\varphi\end{bmatrix}$，$\begin{bmatrix}0\\1\end{bmatrix}$旋转到$\begin{bmatrix}-\sin\varphi\\ \cos\varphi\end{bmatrix}$，见图 1-40. 根据定理 10，

$$\boldsymbol{A}=\begin{bmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{bmatrix}$$

1.8 节的例题 5 是这个变换取 $\varphi=\pi/2$ 时的特例.

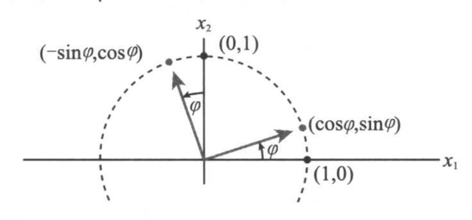


图 1-40　旋转变换

■

1.9.1　R^2 中几何的线性变换

例题 2 和例题 3 都是有几何表示的线性变换的例子. 表 1-2 至表 1-5 列出其他一些常用的平面线性变换. 由于这些变换都是线性的，因此完全由它们在 $\boldsymbol{I}_2$ 的列上的作用所确定. 表中不仅标出 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 的像，而且绘出了变换在整个单位正方形上的作用(见图 1-41).

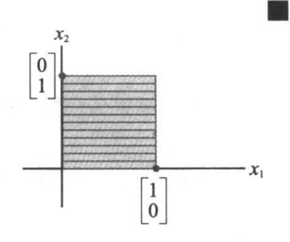
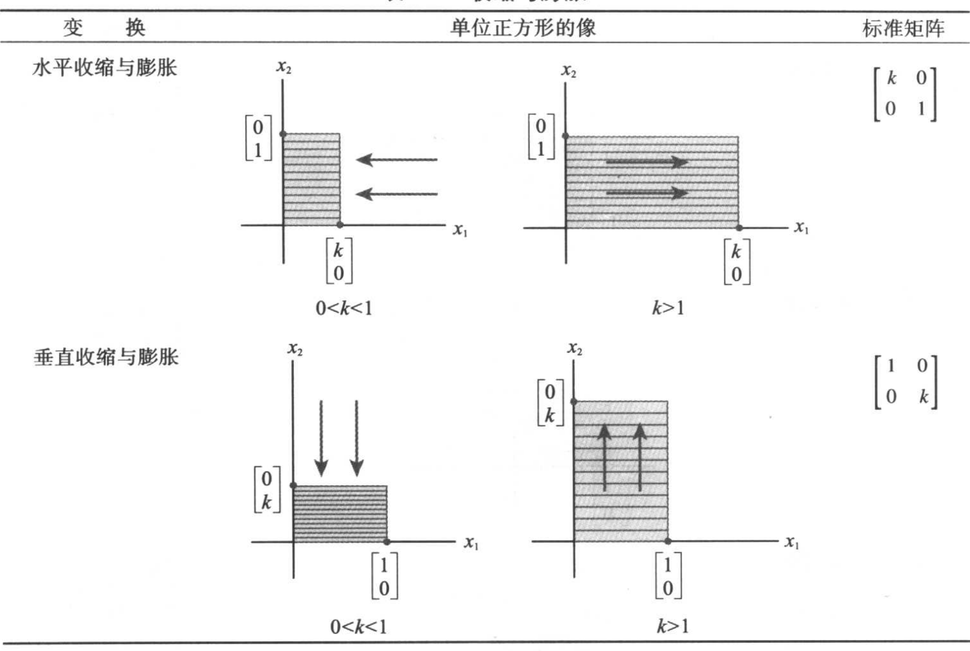


图 1-41　单位正方形

表 1-2 反射

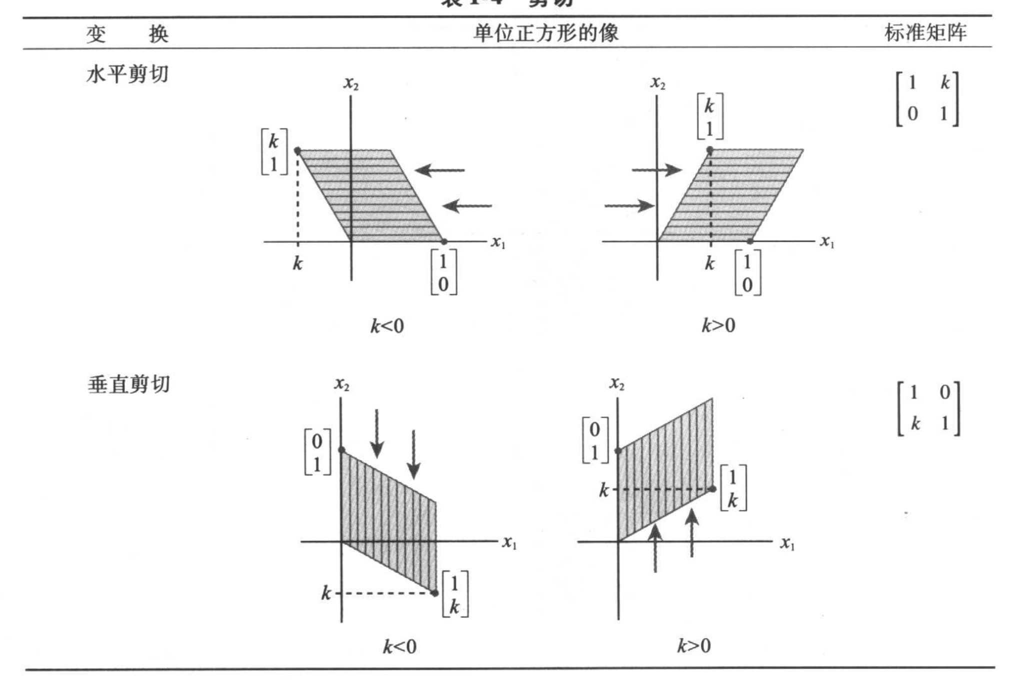
变换	单位正方形的像	标准矩阵
关于 x_1 轴的反射	x_2, $\begin{bmatrix}1\\0\end{bmatrix}$, x_1, $\begin{bmatrix}0\\-1\end{bmatrix}$	$\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$
关于 x_2 轴的反射	x_2, $\begin{bmatrix}0\\1\end{bmatrix}$, x_1, $\begin{bmatrix}-1\\0\end{bmatrix}$	$\begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix}$
关于 $x_1=x_2$ 的反射	x_2, $x_2=x_1$, $\begin{bmatrix}0\\1\end{bmatrix}$, x_1, $\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$
关于 $x_2=-x_1$ 的反射	x_2, $\begin{bmatrix}-1\\0\end{bmatrix}$, x_1, $x_2=-x_1$, $\begin{bmatrix}0\\-1\end{bmatrix}$	$\begin{bmatrix}0 & -1\\-1 & 0\end{bmatrix}$
关于原点的反射	x_2, $\begin{bmatrix}-1\\0\end{bmatrix}$, x_1, $\begin{bmatrix}0\\-1\end{bmatrix}$	$\begin{bmatrix}-1 & 0\\0 & -1\end{bmatrix}$

表 1-3　收缩与膨胀



变　换	单位正方形的像	标准矩阵
水平收缩与膨胀	$0<k<1$　　$k>1$	$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$
垂直收缩与膨胀	$0<k<1$　　$k>1$	$\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$

表 1-4　剪切



变　换	单位正方形的像	标准矩阵
水平剪切	$k<0$　　$k>0$	$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$
垂直剪切	$k<0$　　$k>0$	$\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$

表 1-5 投影

变换	单位正方形的像	标准矩阵
到 x_1 轴上的投影	x_2, x_1, $\begin{bmatrix}0\\0\end{bmatrix}$, $\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}$
到 x_2 轴上的投影	x_2, x_1, $\begin{bmatrix}0\\1\end{bmatrix}$, $\begin{bmatrix}0\\0\end{bmatrix}$	$\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$

其他变换可以通过表 1-2 至表 1-5 中的变换叠加而得到. 例如，可以先做水平剪切变换，再做关于 x_2轴的反射. 2.1 节中将证明线性变换的合成也是线性的(也可参见习题 36).

1.9.2 存在性和唯一性问题

线性变换的概念为我们理解存在性和唯一性问题提供了新的方法. 下面介绍变换的两个概念.

【定义】 称映射 T: $\mathbf{R}^n \to \mathbf{R}^m$ 为 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 上($\mathbf{R}^n$ onto $\mathbf{R}^m$)的映射(或满射)，如果 $\mathbf{R}^m$ 中的任意 $\boldsymbol{b}$ 都是 $\mathbf{R}^n$ 中至少一个 $\boldsymbol{x}$ 的像.

等价地，当 T 的值域是整个上域 $\mathbf{R}^m$ 时，T 是到 $\mathbf{R}^m$ 上的映射. 即，如果对上域 $\mathbf{R}^m$ 中的任意 $\boldsymbol{b}$，$T(\boldsymbol{x}) = \boldsymbol{b}$ 都至少存在一个解，则 T 把 $\mathbf{R}^n$ 映射到 $\mathbf{R}^m$ 上. "T 是否是 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 上的映射?" 是个存在性问题. 如果 $\mathbf{R}^m$ 中存在 $\boldsymbol{b}$ 使得方程 $T(\boldsymbol{x}) = \boldsymbol{b}$ 无解，则 T 不是到上的映射. 见图 1-42.

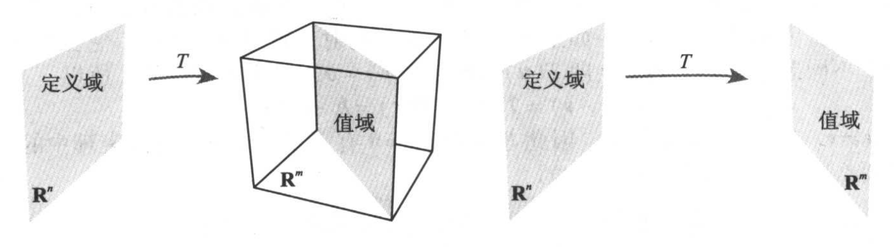


图 1-42 T 的值域是否为整个 $\mathbf{R}^m$

84 ~ 87

【定义】 称映射 $T: \mathbf{R}^n \to \mathbf{R}^m$ 为**一对一**(one-to-one)的映射(或单射)，如果 $\mathbf{R}^m$ 中的任意 $\boldsymbol{b}$ 都是 $\mathbf{R}^n$ 中至多一个 $\boldsymbol{x}$ 的像.

等价地，如果对 $\mathbf{R}^m$ 中的任意 $\boldsymbol{b}$，方程 $T(\boldsymbol{x})=\boldsymbol{b}$ 有唯一解或者无解，则 T 为一对一映射. "T 是否是一对一映射?"是个唯一性问题. 如果存在 $\mathbf{R}^m$ 中的 $\boldsymbol{b}$，$\boldsymbol{b}$ 是 $\mathbf{R}^n$ 中多于一个向量的像，则 T 不是一对一映射. 如果不存在这样的 $\boldsymbol{b}$，T 就是一对一映射. 见图 1-43.

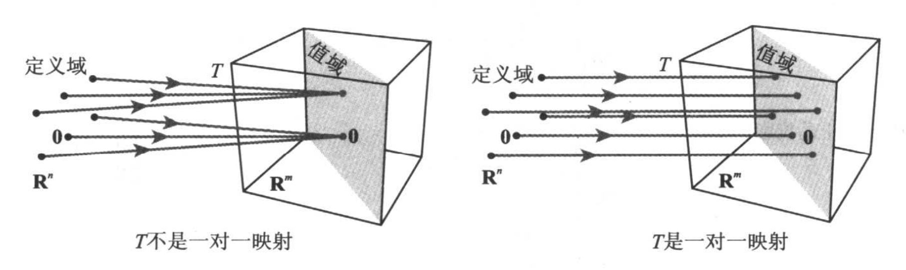


图 1-43　$\boldsymbol{b}$ 是否是至多一个向量的像

表 1-5 中所示的投影映射不是一对一映射，并且也不是从 $\mathbf{R}^n$ 到 $\mathbf{R}^n$ 上的映射. 表 1-2，表 1-3，表 1-4 中的变换都是一对一映射，并且也都是从 $\mathbf{R}^n$ 到 $\mathbf{R}^n$ 上的映射. 其他情形将在下面两道例题中给出.

【例题 4】 设 T 为一个线性变换，其标准矩阵为

$$\boldsymbol{A}=\begin{bmatrix}1 & -4 & 8 & 1\\ 0 & 2 & -1 & 3\\ 0 & 0 & 0 & 5\end{bmatrix}$$

T 是否是从 $\mathbf{R}^4$ 到 $\mathbf{R}^3$ 上的映射? T 是否是一对一映射?

解: 由于 $\boldsymbol{A}$ 是阶梯形式，我们可以立刻看出 $\boldsymbol{A}$ 的每一行中都有一个主元位置. 根据 1.4 节定理 4，对 $\mathbf{R}^n$ 中的任意 $\boldsymbol{b}$，方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 都相容，即线性变换 T 把 $\mathbf{R}^4$ 映射到 $\mathbf{R}^3$ 上. 然而，由于方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 含有自由变量(方程有四个变量，只有 3 个是基本变量)，每个 $\boldsymbol{b}$ 都是不止一个 $\boldsymbol{x}$ 的像，即 T 不是一对一映射. ■

【定理 11】 设 $T: \mathbf{R}^n \to \mathbf{R}^m$ 为线性变换. 则 T 为一对一映射当且仅当方程 $T(\boldsymbol{x})=\mathbf{0}$ 仅有平凡解.

证明: 由于 T 是线性的，则 $T(\mathbf{0})=\mathbf{0}$. 如果 T 为一对一映射，则方程 $T(\boldsymbol{x})=\mathbf{0}$ 至多有一个解，因此仅有平凡解. 如果 T 不是一对一映射，则存在一个 $\boldsymbol{b}$，它至少是 $\mathbf{R}^n$ 中两个不同向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的像. 即 $T(\boldsymbol{u})=\boldsymbol{b}$ 且 $T(\boldsymbol{v})=\boldsymbol{b}$. 但 T 是线性的，所以

88

$$T(\boldsymbol{u}-\boldsymbol{v})=T(\boldsymbol{u})-T(\boldsymbol{v})=\boldsymbol{b}-\boldsymbol{b}=\mathbf{0}$$

由于 $\boldsymbol{u}\neq\boldsymbol{v}$，向量 $\boldsymbol{u}-\boldsymbol{v}$ 不为零. 因此方程 $T(\boldsymbol{x})=\mathbf{0}$ 有不止一个解. 所以定理中的两种情况要么同时成立要么同时不成立. ■

【定理 12】 设 $T: \mathbf{R}^n \to \mathbf{R}^m$ 为一个线性变换，且 $\boldsymbol{A}$ 为 T 的标准矩阵. 则:

(a) T 是 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 上的映射当且仅当 $\boldsymbol{A}$ 中的列张成 $\mathbf{R}^m$.

(b) T 是一对一映射当且仅当 $\boldsymbol{A}$ 中的列线性无关.

证明：

(a)根据1.4节定理4，$\boldsymbol{A}$ 中的列张成 $\mathbf{R}^m$ 当且仅当对任意 $\boldsymbol{b}$，方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 相容，即当且仅当对任意 $\boldsymbol{b}$，方程 $T(\boldsymbol{x})=\boldsymbol{b}$ 至少有一个解. 而最后这一命题又当且仅当 T 是 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 上的映射.

(b)方程 $T(\boldsymbol{x})=\mathbf{0}$ 和 $\boldsymbol{Ax}=\mathbf{0}$ 含义相同,只是记号不同. 因此根据定理11, T 是一对一映射当且仅当 $\boldsymbol{Ax}=\mathbf{0}$ 仅有平凡解. 而根据1.7.1节的命题(3)可知，$\boldsymbol{Ax}=\mathbf{0}$ 仅有平凡解当且仅当 $\boldsymbol{A}$ 中的列线性无关. ■

定理12中的命题(a)等价于：T 是 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 上的映射当且仅当 $\mathbf{R}^m$ 中的任意向量都是 $\boldsymbol{A}$ 中列的线性组合. 见1.4节定理4.

在下面的例题以及本节部分习题中，列向量都写成行的形式，例如 $\boldsymbol{x}=(x_1,x_2)$，此外 $T(\boldsymbol{x})$ 的正式写法是 $T((x_1,x_2))$，不过我们简记为 $T(x_1,x_2)$.

【例题5】 设 $T(x_1,x_2)=T(3x_1+x_2,5x_1+7x_2,x_1+3x_2)$. 证明 T 是一对一的线性映射. T 是从 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 上的映射吗？

解： 令 T 的标准矩阵为 $\boldsymbol{A}$，将 $\boldsymbol{x}$ 和 $T(\boldsymbol{x})$ 写成列向量的形式，通过观察 $\boldsymbol{Ax}$ 中元素的行向量计算可以确定 $\boldsymbol{A}$.

$$T(\boldsymbol{x})=\begin{bmatrix}3x_1+x_2\\5x_1+7x_2\\x_1+3x_2\end{bmatrix}=\underbrace{\begin{bmatrix}?&?\\?&?\\?&?\end{bmatrix}}_{\boldsymbol{A}}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}3&1\\5&7\\1&3\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}\quad(4)$$

因此 T 的确是一个线性变换，它的标准矩阵如(4)中所示. 由于 $\boldsymbol{A}$ 的各列没有倍数关系，$\boldsymbol{A}$ 的列线性无关. 根据定理12(b)，T 是一对一映射，为确定 T 是从 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 上的映射，考察 $\boldsymbol{A}$ 的列所张成的子空间. 由于 $\boldsymbol{A}$ 为 3×2 矩阵，根据定理4，$\boldsymbol{A}$ 的列张成 $\mathbf{R}^3$ 当且仅当 $\boldsymbol{A}$ 有3个主元位置. 但这是不可能的，因为 $\boldsymbol{A}$ 仅有2列. 因此 $\boldsymbol{A}$ 中的列不能张成 $\mathbf{R}^3$，且对应的线性变换不是到 $\mathbf{R}^3$ 上的映射.（见图1-44） ■

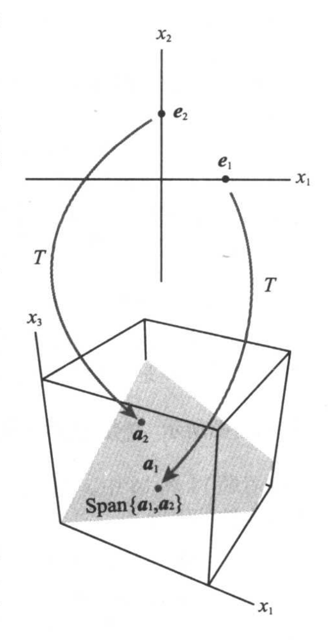


图1-44 变换 T 不是到 $\mathbf{R}^3$ 上的映射

89

基础练习

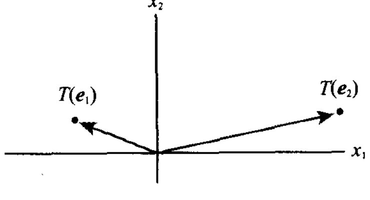
设 T: $\mathbf{R}^2\to\mathbf{R}^2$ 为一个变换，它首先执行把 $\boldsymbol{e}_2$ 映射到 $\boldsymbol{e}_2-0.5\boldsymbol{e}_1$ 的水平剪切(保持 $\boldsymbol{e}_1$ 不变)，然后将该结果做关于 $\boldsymbol{x}_2$ 轴的反射. 假设 T 是线性的，求其标准矩阵. [提示：确定 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 像的位置.]

习题 1.9

在习题1~10中，假设 T 为线性变换. 求出 T 的标准矩阵.

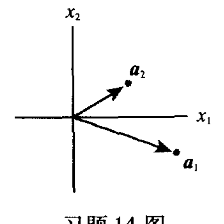
1. T: $\mathbf{R}^2\to\mathbf{R}^4$，$T(\boldsymbol{e}_1)=(3,1,3,1)$，$T(\boldsymbol{e}_2)=(-5,2,0,0)$. 其中 $\boldsymbol{e}_1=(1,0)$，$\boldsymbol{e}_2=(0,1)$.
2. T: $\mathbf{R}^3\to\mathbf{R}^2$，$T(\boldsymbol{e}_1)=(1,3)$，$T(\boldsymbol{e}_2)=(4,-7)$. $T(\boldsymbol{e}_3)=(-5,4)$. 其中 $\boldsymbol{e}_1$，$\boldsymbol{e}_2$，$\boldsymbol{e}_3$ 为 3×3 单位矩阵中的列.

3. T: $\mathbf{R}^2 \to \mathbf{R}^2$ 把点绕原点逆时针旋转了 $3\pi/2$ 角度.
4. T: $\mathbf{R}^2 \to \mathbf{R}^2$ 把点顺时针旋转了 $-\pi/4$ 角度. [提示: $T(\boldsymbol{e}_1)=(1/\sqrt{2}, -1/\sqrt{2})$]
5. T: $\mathbf{R}^2 \to \mathbf{R}^2$ 为垂直剪切变换, 把 $\boldsymbol{e}_1$ 映射到 $\boldsymbol{e}_1 - 2\boldsymbol{e}_2$, 且保持 $\boldsymbol{e}_2$ 不变.
6. T: $\mathbf{R}^2 \to \mathbf{R}^2$ 为水平剪切变换, 把 $\boldsymbol{e}_2$ 映射到 $\boldsymbol{e}_2 + 3\boldsymbol{e}_1$, 且保持 $\boldsymbol{e}_1$ 不变.
7. T: $\mathbf{R}^2 \to \mathbf{R}^2$ 首先把点顺时针旋转 $-3\pi/4$ 角度, 然后再做关于横轴 $\boldsymbol{x}_1$ 的反射. [提示: $T(\boldsymbol{e}_1) = (-1/\sqrt{2}, 1/\sqrt{2})$]
8. T: $\mathbf{R}^2 \to \mathbf{R}^2$ 首先将点做关于横轴 $\boldsymbol{x}_1$ 的反射. 然后再做关于直线 $\boldsymbol{x}_2 = \boldsymbol{x}_1$ 的反射.
9. T: $\mathbf{R}^2 \to \mathbf{R}^2$ 首先进行水平剪切, 把 $\boldsymbol{e}_2$ 映射到 $\boldsymbol{e}_2 - 2\boldsymbol{e}_1$(保持 $\boldsymbol{e}_1$ 不变), 然后再做关于直线 $\boldsymbol{x}_2 = -\boldsymbol{x}_1$ 的反射.
10. T: $\mathbf{R}^2 \to \mathbf{R}^2$ 首先做关于竖轴 $\boldsymbol{x}_2$ 的反射, 然后把点旋转 $\pi/2$ 角度.
11. 线性变换 T: $\mathbf{R}^2 \to \mathbf{R}^2$ 首先做关于 $\boldsymbol{x}_1$ 轴的反射, 然后再做关于 $\boldsymbol{x}_2$ 轴的反射. 证明 T 可以表示成一个绕原点旋转的线性变换. 旋转的角度是多少?
12. 证明习题 8 中的变换是绕原点的旋转. 旋转的角度是多少?
13. 设 T: $\mathbf{R}^2 \to \mathbf{R}^2$ 是一个线性变换, 且 $T(\boldsymbol{e}_1)$ 和 $T(\boldsymbol{e}_2)$ 如右图所示. 利用图示, 画出向量 $T(2,1)$.



习题 13 图

14. 设 T: $\mathbf{R}^2 \to \mathbf{R}^2$ 是一个线性变换, 且标准矩阵为 $A = [\boldsymbol{a}_1 \quad \boldsymbol{a}_2]$, $\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2$ 如右图所示. 利用图示, 画出 $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$ 在变换 T 下的像.



习题 14 图

在习题 15 和 16 中, 假设方程对变量的所有取值都成立, 填补矩阵中空缺的元素.

15. $\begin{bmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3x_1 - \quad 2x_3 \\ 4x_1 \\ x_1 - x_2 + x_3 \end{bmatrix}$　16. $\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ -2x_1 + x_2 \\ x_1 \end{bmatrix}$

90

在习题 17 ~ 20 中, 通过找出构成映射的矩阵来证明 T 是线性变换. 注意 $x_1, x_2, \cdots$ 不是向量, 而是向量中的元素.

17. $T(x_1, x_2, x_3, x_4) = (0, x_1 + x_2, x_2 + x_3, x_3 + x_4)$
18. $T(x_1, x_2) = (2x_2 - 3x_1, x_1 - 4x_2, 0, x_2)$
19. $T(x_1, x_2, x_3) = (x_1 - 5x_2 + 4x_3, x_2 - 6x_3)$
20. $T(x_1, x_2, x_3, x_4) = 2x_1 + 3x_3 - 4x_4 \quad (T: \mathbf{R}^4 \to \mathbf{R})$
21. 设 T: $\mathbf{R}^2 \to \mathbf{R}^2$ 为线性变换, 且 $T(\boldsymbol{x}_1, \boldsymbol{x}_2) = (\boldsymbol{x}_1 + \boldsymbol{x}_2, 4\boldsymbol{x}_1 + 5\boldsymbol{x}_2)$. 求出使得 $T(\boldsymbol{x}) = (3, 8)$ 的 $\boldsymbol{x}$.
22. 设 T: $\mathbf{R}^2 \to \mathbf{R}^3$ 为线性变换, 且 $T(\boldsymbol{x}_1, \boldsymbol{x}_2) = (\boldsymbol{x}_1 - 2\boldsymbol{x}_2, -\boldsymbol{x}_1 + 3\boldsymbol{x}_2, 3\boldsymbol{x}_1 - 2\boldsymbol{x}_2)$. 求出使 $T(\boldsymbol{x}) = (-1, 4, 9)$ 的 $\boldsymbol{x}$.

在习题 23 和 24 中, 判断各命题的真假, 并说明理由.

23. a. 一个线性变换 T: $\mathbf{R}^n \to \mathbf{R}^n$ 完全由它在 $n \times n$ 单位矩阵的列上的作用所确定.
    b. 如果变换 T: $\mathbf{R}^2 \to \mathbf{R}^2$ 把向量绕原点旋转 φ 角度, 则变换 T 是线性的.
    c. 当两个线性变换相继执行, 它们的合成不一定是线性的.
    d. 一个映射 T: $\mathbf{R}^n \to \mathbf{R}^m$, 如果它把 $\mathbf{R}^n$ 中的每个向量 $\boldsymbol{x}$ 映射到 $\mathbf{R}^m$ 中的一些向量上, 则 T 是

到 $\mathbf{R}^m$ 上的映射.

e. 如果 A 为 3×2 矩阵，则变换 $x\mapsto Ax$ 不可能是一对一映射.

24. a. 不是每一个从 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 上的线性变换都是矩阵变换.

b. A 是从 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 上线性变换的标准矩阵，A 中的列是 $n\times n$ 单位矩阵中列的像.

c. 从 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 上将点做关于横轴、竖轴或原点反射的线性变换的标准矩阵形如 $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$，这里 a 和 b 为 ± 1.

d. 如果一个映射 T：$\mathbf{R}^n\to\mathbf{R}^m$ 把 $\mathbf{R}^n$ 中的每个向量 x 映射到 $\mathbf{R}^m$ 中唯一一个向量上，则它是一对一映射.

e. 如果 A 为 3×2 矩阵，则变换 $x\mapsto Ax$ 不可能把 $\mathbf{R}^2$ 映射到 $\mathbf{R}^3$ 上.

在习题 25 ~ 28 中，判断已知的线性变换是(a)一对一映射，还是(b)到上映射. 给出证明.

25. 习题 17 中的变换.
26. 习题 2 中的变换.
27. 习题 19 中的变换.
28. 习题 14 中的变换.

在习题 29 和 30 中，写出线性变换 T 的标准矩阵可能的阶梯型式. 使用 1.2 节例题 1 的符号.

29. T：$\mathbf{R}^3\to\mathbf{R}^4$ 为一对一映射.
30. T：$\mathbf{R}^4\to\mathbf{R}^3$ 为到上映射.
31. 设 T：$\mathbf{R}^n\to\mathbf{R}^m$ 为一个线性变换，A 为标准矩阵. 填补命题使其正确："T 是一对一映射当且仅当 A 有________个主元列." 并给出解释. [提示：见 1.7 节后的习题.]
32. 设 T：$\mathbf{R}^n\to\mathbf{R}^m$ 为一个线性变换，A 为标准矩阵. 填补命题使其正确："T 把 $\mathbf{R}^n$ 映射到 $\mathbf{R}^m$ 上当且仅当 A 有________个主元列." 并找出相应的定理给予解释.
33. 改写定理 10 中 A 的唯一性：设 T：$\mathbf{R}^n\to\mathbf{R}^m$ 为一个线性变换，且 $T(x)=Bx$，B 为 $m\times n$ 矩阵. 如果 A 为 T 的标准矩阵，则 $A=B$. [提示：证明 A 和 B 中有相同的列.]
34. 为什么"线性变换 T 是否是到上的?"是一个存在性问题?
35. 如果一个线性变换 T：$\mathbf{R}^n\to\mathbf{R}^m$ 把 $\mathbf{R}^n$ 映射到 $\mathbf{R}^m$ 上，n 与 m 有什么关系? 如果 T 是一对一的，n 与 m 又有什么关系?
36. 设 S：$\mathbf{R}^p\to\mathbf{R}^n$ 和 T：$\mathbf{R}^n\to\mathbf{R}^m$ 均为线性变换. 证明映射 $x\mapsto T(S(x))$ 是一个线性变换(从 $\mathbf{R}^p$ 到 $\mathbf{R}^m$ 上). [提示：对 $\mathbf{R}^p$ 中的 u 和 v，以及数量 c 和 d，计算 $T(S(cu+dv))$. 对每一步计算进行验证，并解释为什么该计算能证明结论.]

[M]在习题 37 ~ 40 中，设 T 为线性变换，它的标准矩阵如题给出. 判断习题 37 和 38 中的 T 是否为一对一映射；习题 39 和 40 中的 T 是否为从 $\mathbf{R}^5$ 到 $\mathbf{R}^5$ 上的映射. 给出证明.

37. $\begin{bmatrix} -5 & 10 & -5 & 4 \\ 8 & 3 & -4 & 7 \\ 4 & -9 & 5 & -3 \\ -3 & -2 & 5 & 4 \end{bmatrix}$

38. $\begin{bmatrix} 7 & 5 & 4 & -9 \\ 10 & 6 & 16 & -4 \\ 12 & 8 & 12 & 7 \\ -8 & -6 & -2 & 5 \end{bmatrix}$

39. $\begin{bmatrix} 4 & -7 & 3 & 7 & 5 \\ 6 & -8 & 5 & 12 & -8 \\ -7 & 10 & -8 & -9 & 14 \\ 3 & -5 & 4 & 2 & -6 \\ -5 & 6 & -6 & -7 & 3 \end{bmatrix}$

40. $\begin{bmatrix} 9 & 13 & 5 & 6 & -1 \\ 14 & 15 & -7 & -6 & 4 \\ -8 & -9 & 12 & -5 & -9 \\ -5 & -6 & -8 & 9 & 8 \\ 13 & 14 & 15 & 2 & 11 \end{bmatrix}$

基础练习答案

根据 T 在 e_1 和 e_2 上的作用解题(见图 1-45). 首先，e_1 在剪切变换下没有变化，然后反射到 $-e_1$，因此 $T(e_1)=-e_1$. 第 2 步,e_2 通过剪切变换得到 $e_2-0.5e_1$. 关于 x_2 轴的反射把 e_1 变成 $-e_1$,且 e_2 保持不变,向量 $e_2-0.5e_1$ 变成 $e_2+0.5e_1$. 所以 $T(e_2)=e_2+0.5e_1$. 因此 T 的标准矩阵是

$$[T(e_1) \quad T(e_2)]=[-e_1 \quad e_2+0.5e_1]=\begin{bmatrix}-1 & 0.5\\ 0 & 1\end{bmatrix}$$

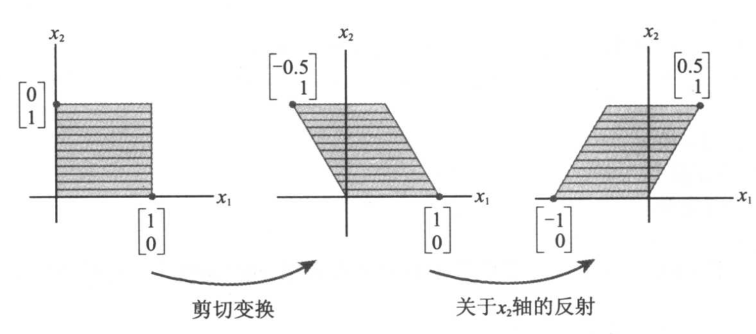


图 1-45 变换的合成

1.10 商业、科学和工程学中的线性模型

本节中的数学模型都是线性的，即每个模型都用线性方程组来表示，通常写成向量或矩阵的形式. 第一个模型与营养学有关，它实际上代表了线性规划问题的一类通用技术. 第二个模型源于电气工程学. 第三个模型引入了**线性差分方程**的概念，这是研究许多领域内动态过程的一种强有力的工具，包括工程学、生态学、经济学、电信学以及管理科学. 由于自然现象通常都是线性的，或者当变量取值在合理范围内时近似于线性，因此线性模型非常重要. 此外，线性模型比复杂的非线性模型更易于用计算机进行计算.

92

学习下面每个模型时，请注意其线性性是如何反映在建模系统当中的.

1.10.1 制作有营养的减肥食谱

剑桥食谱流行于 20 世纪 80 年代，其配方建立在多年的科学研究基础上. 由 A. H. 霍华德博士领导的科研团队经过对肥胖病人进行八年多的临床试验后，在剑桥大学提出了这份食谱. [1] 在这个低卡路里的粉末状食物配方中，碳水化合物、高蛋白质、脂肪，连同维生素、矿物质、微量元素和电解质达到了精确的平衡. 数百万人使用这份食谱快速且有效地减轻了体重.

为了得到预期的营养比例以及总量，霍华德博士需要在食谱中混合大量的食物. 每一种食物都提供了一些所需的成分，但比例可能不正确. 例如，脱脂牛奶是蛋白

---

1. 这个快速减肥法首次发表于 *International Journal of Obesity* (1978) **2**, 321 - 332.

质的主要来源，但它含有太多的钙. 因此使用含钙很少的大豆粉作为蛋白质的补充来源. 然而，大豆粉中脂肪含量过高，因此加入脂肪含量少于钙的乳清. 但乳清中却又含有过多的碳水化合物……

下面例题的规模相对较小，仅仅考虑了少数几种食物以及营养. 表1-6列出了食谱中的三种成分，以及每100克(g)成分所提供的各种营养的量.[1]

表　1-6

营　养	每100克成分(g)所含的营养			剑桥食谱一天所提供的营养
	脱脂牛奶	大豆粉	乳清	
蛋白质	36	51	13	33
碳水化合物	52	34	74	45
脂肪	0	7	1.1	3

【例题1】　试求脱脂牛奶、大豆粉和乳清的某种组合，使得混合物提供的蛋白质、碳水化合物和脂肪如表1-6中所示.

解： 设x_1，x_2和x_3分别表示这三种食物的量(单位：100g). 求解本题的一个方法是列出关于每一种营养的方程. 例如，

$$\{x_1\text{单位脱脂牛奶}\}\cdot\{1\text{单位脱脂牛奶中蛋白质的含量}\}$$

93

这个乘积给出了x_1单位的脱脂牛奶提供的蛋白质总量. 另外，我们可以加上大豆粉和乳清的类似的乘积，令上述乘积之和等于所需蛋白质的总量. 对于每一种营养都可以作类似的计算.

还有一个更有效的方法，其概念亦十分简单，即对每一种食物考虑一个“营养向量”，并建立一个方程. x_1单位脱脂牛奶提供的营养总量是下列数乘：

$$\underset{\text{数量}}{\{x_1\text{单位脱脂牛奶}\}}\cdot\underset{\text{向量}}{\{1\text{单位脱脂牛奶的营养向量}\}}=x_1\boldsymbol{a}_1 \tag{1}$$

其中$\boldsymbol{a}_1$是表1-6中的第1列. 设$\boldsymbol{a}_2$和$\boldsymbol{a}_3$分别为大豆粉和乳清对应的营养向量，并且设向量$\boldsymbol{b}$列出了所需的营养总量(表1-6中的最后一列). 则$x_2\boldsymbol{a}_2$和$x_3\boldsymbol{a}_3$分别给出了x_2单位大豆粉和x_3单位乳清所提供的营养. 所以我们需要的方程为

$$x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+x_3\boldsymbol{a}_3=\boldsymbol{b} \tag{2}$$

行化简其对应方程组的增广矩阵，得到

$$\begin{bmatrix}36 & 51 & 13 & 33\\52 & 34 & 74 & 45\\0 & 7 & 1.1 & 3\end{bmatrix}\sim\cdots\sim\begin{bmatrix}1 & 0 & 0 & 0.277\\0 & 1 & 0 & 0.392\\0 & 0 & 1 & 0.233\end{bmatrix}$$

为了提供所所需的蛋白质、碳水化合物和脂肪总量，食谱中需要包含0.227单位的脱脂牛奶，0.392单位的大豆粉，0.233单位的乳清(保留三位有效数字). ■

上面求出的x_1，x_2，x_3的值都非负，这是解确实可行的必要条件(否则，我们怎样配制-0.233单位的乳清?). 当配制一份营养丰富的食物时，需要使用大量的食物，然后求对应方程组的“非负”解. 因此，为了找出有非负解的方程组，需要检

1. 1984年食谱中的成分；成分中的营养数据修改自USDA Agricultural Handbooks No. 8－1和8－6，1976.

验大量不同食物成分的组合．事实上，剑桥食谱的制造商可以仅用33种成分来精确提供31种营养．

这个食谱构造问题引出了线性方程(2)，因为如(1)所示，每种食品提供的营养总量都可以写成向量数乘的形式．即，一种食物提供的营养与添加到混合食谱中的该食物的量成正比．此外，混合物中每种营养都是各食物成分中该营养的总和．

实际生活中经常需要为人或牲畜配制特别的食谱．它们通常需要用到线性规划技术，这里所介绍的构造向量方程的方法往往可以简化其求解．

94

1.10.2　线性方程组与电网

简单电网中的电流可以用线性方程组来描述，电压电源(例如电池)迫使电子在电网中流动形成电流．当电流经过电阻(例如灯泡或者发动机)时，一些电压被“消耗”．根据欧姆定律，流经电阻时的“电压降”由下列公式给出：

$$V = RI$$

电压 V、电阻 R 和电流 I 分别以伏特、欧姆(记作(Ω))和安培为单位．

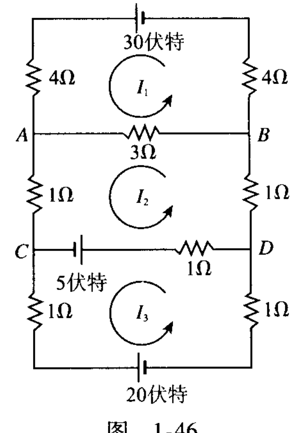


图　1-46

图1-46中的电网连接了三个闭回路．回路1，2和3中的电流分别用 I_1，I_2 和 I_3 表示．回路电流的方向是任意的．如果一个电流为负，则表示实际的电流方向与图中闭回路的电流方向相反．如果电流所示的方向由电池(╢├)正极(长的一端)指向负极(短的一端)，则电压为正；否则电压为负．

回路中的电流服从下列定律：

基尔霍夫电压定律

沿某个方向环绕回路一周的所有电压降 RI 的代数和等于沿同一方向环绕该回路一周的电源电压的代数和．

【例题2】　确定图1-46电网中的回路电流．

解：回路1中，电流 I_1 流过三个电阻，且电压降 RI 为：

$$4I_1 + 4I_1 + 3I_1 = (4 + 4 + 3)I_1 = 11I_1$$

回路2中的电流也流经回路1的一部分，即从A到B的分支，对应的电压降 RI 为 $3I_2$ 伏特．然而，回路1中电流在AB段的方向与回路2中选定的方向相反，因此回路1中所有电压降 RI 的代数和为 $11I_1 - 3I_2$．由于回路1中的电压为+30伏特，基尔霍夫电压定律表明

$$11I_1 - 3I_2 = 30$$

回路2的方程为

$$-3I_1 + 6I_2 - I_3 = 5$$

其中，$-3I_1$ 是回路1中流经AB分支的电流(因为电流与回路2中的电流方向相反，所以电压为负)；$6I_2$ 是回路2中所有的电阻乘上回路电流的和；$-I_3 = -1 \cdot I_3$ 是回路

95

3中流经CD分支上1欧姆电阻的电流，方向与回路2中该段的电流方向相反．回路

3 的方程为

$$-I_2+3I_3=-25$$

注意，在 CD 分支上 5 伏特的电池被当作是回路 2 和回路 3 中的一部分，但是由于回路 3 中电流方向，电池在回路 3 中为 −5 伏特. 出于同样的道理，20 伏特的电池也应取负值.

回路电流要通过求解下列方程组得到

$$\begin{aligned} 11I_1-3I_2 &= 30 \\ -3I_1+6I_2-\ I_3 &= 5 \\ -\ I_2+3I_3 &= -25 \end{aligned} \tag{3}$$

对增广矩阵进行行变换，得到解：$I_1=3$ 安培，$I_2=1$ 安培，$I_3=-8$ 安培. I_3取负值说明回路 3 中的实际电流与图 1-46 中显示的电流方向相反. ■

把方程组(3)看成向量方程将很有启发：

$$I_1\begin{bmatrix}11\\-3\\0\end{bmatrix}+I_2\begin{bmatrix}-3\\6\\-1\end{bmatrix}+I_3\begin{bmatrix}0\\-1\\3\end{bmatrix}=\begin{bmatrix}30\\5\\-25\end{bmatrix} \tag{4}$$

$$\uparrow\ \boldsymbol{r}_1 \qquad \uparrow\ \boldsymbol{r}_2 \qquad \uparrow\ \boldsymbol{r}_3 \qquad \uparrow\ \boldsymbol{v}$$

每个向量中的第一个元素都与回路 1 有关，第二、三个元素也类似. 第一个电阻向量 $\boldsymbol{r}_1$列出了电流 I_1流经的不同回路中的电阻，并且当电流 I_1的方向与某个回路中的电流方向相反时，我们将电阻取负值. 观察图 1-46，看 $\boldsymbol{r}_1$中的元素是如何计算出来的；然后再考察 $\boldsymbol{r}_2$和 $\boldsymbol{r}_3$. (4)的矩阵形式为

$$R\boldsymbol{i}=\boldsymbol{v},\text{其中 } R=[\boldsymbol{r}_1\quad \boldsymbol{r}_2\quad \boldsymbol{r}_3],\boldsymbol{i}=\begin{bmatrix}I_1\\I_2\\I_3\end{bmatrix}$$

它就是用矩阵表示的欧姆定律. 如果所有的回路电流都选定同一方向(比如逆时针)，则 R 的主对角线以外的所有元素都为负.

容易看出，矩阵方程 $R\boldsymbol{i}=\boldsymbol{v}$ 保证了该模型是线性的. 即，如果电压向量变成原来的两倍，则电流向量必然是原来的两倍. 叠加原理同样成立，即方程(4)的解是下列方程解的和.

$$R\boldsymbol{i}=\begin{bmatrix}30\\0\\0\end{bmatrix},\quad R\boldsymbol{i}=\begin{bmatrix}0\\5\\0\end{bmatrix},\quad R\boldsymbol{i}=\begin{bmatrix}0\\0\\-25\end{bmatrix}$$

上述每个方程都对应于仅含一个电压电源的电路(其余电源都用电线代替). 电流的模型是线性的，这是因为欧姆定律和基尔霍夫电压定律恰好是线性的：经过一个电阻的电压降与流经它的电流成正比，且回路中电压降之和与电源电压之和相等(基尔霍夫电压定律).

96

电网中的回路电流可以用来确定电网中每一分支中的电流. 如果只有一个回路电流流经一个分支，例如图 1-46 中的 BD，则分支电流等于回路电流. 如果多于一个

回路电流流经一个分支，例如从 AB，则分支电流为该分支中回路电流的代数和(基尔霍夫电压定律). 例如，AB 分支中的电流为 $I_1-I_2=3-1=2$ 安培，方向与 I_1 相同，CD 分支中的电流为 $I_1+I_2=9$ 安培.

1.10.3　差分方程

生态学、经济学和工程学等许多领域中经常需要对随时间变化的动态系统进行数学建模. 系统中的某些量按离散时间间隔来测量，于是产生了向量序列 $\boldsymbol{x}_0,\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k$ 中的元素给出了第 k 次测量时系统状态的有关信息.

如果存在矩阵 $\boldsymbol{A}$，使得 $\boldsymbol{x}_1=\boldsymbol{A}\boldsymbol{x}_0$，$\boldsymbol{x}_2=\boldsymbol{A}\boldsymbol{x}_1$，依此类推，

$$\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k \qquad k=0,1,2,\cdots \tag{5}$$

则称(5)为一个**线性差分方程**(linear difference equation) 或者**递归方程**(recurrence relation). 给定这样一个方程，如果已知 $\boldsymbol{x}_0$，则可以计算 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 等等. 4.8、4.9 节以及第 5 章的几节将改进 $\boldsymbol{x}_k$ 的计算公式，并且描述当 k 增至无限时 $\boldsymbol{x}_k$ 的变化情况. 下面的讨论举例说明了实际问题中产生的差分方程.

人口统计学家感兴趣的一个题目是人口迁移或人群从一个地区到另一个地区的移动. 这里我们考虑一个简单的模型，即某城市及其周边郊区在若干年内的人口变化.

设定一个初始的年份，比如说 2000 年，用 r_0，r_1 分别表示这一年城市和郊区的人口. 设 $\boldsymbol{x}_0$ 为人口向量.

$$\boldsymbol{x}_0=\begin{bmatrix} r_0 \\ s_0 \end{bmatrix} \qquad \begin{matrix}\text{2000 年城市人口} \\ \text{2000 年郊区人口}\end{matrix}$$

对 2001 年以及后面的年份，都用向量表示出每一年城市和郊区的人口

$$\boldsymbol{x}_1=\begin{bmatrix} r_1 \\ s_1 \end{bmatrix},\quad \boldsymbol{x}_2=\begin{bmatrix} r_2 \\ s_2 \end{bmatrix},\quad \boldsymbol{x}_3=\begin{bmatrix} r_3 \\ s_3 \end{bmatrix},\cdots$$

我们的目标是用数学公式表示出这些向量之间的关系.

假设人口统计学家的研究表明，每年大约有 5% 的城市人口迁移到郊区(95% 仍然留在城市)，有 3% 的郊区人口迁移到城市(97% 仍然留在郊区). 见图 1-47.

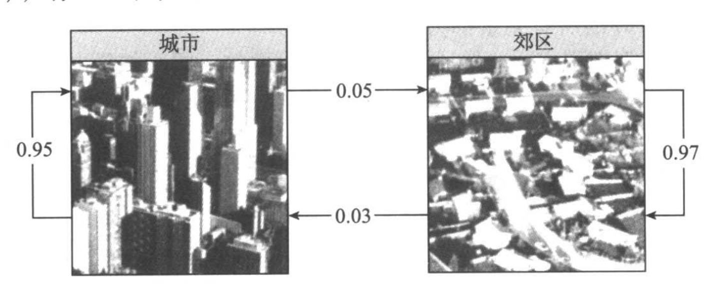


图 1-47　城市和郊区每年的迁移人口百分比

97

一年之后，城市的初始人口 r_0 在城市和郊区的分布为:

$$\begin{bmatrix} 0.95r_0 \\ 0.05r_0 \end{bmatrix}=r_0\begin{bmatrix} 0.95 \\ 0.05 \end{bmatrix} \qquad \begin{matrix}\text{留在城市} \\ \text{移居郊区}\end{matrix} \tag{6}$$

一年之后，2000 年郊区人口 s_0的分布为：

$$s_0\begin{bmatrix}0.03\\0.97\end{bmatrix}\quad\begin{matrix}\text{移居城市}\\\text{留在郊区}\end{matrix}\tag{7}$$

(6)和(7)中的向量解释了2001 年全部人口的分布.[1]因此

$$\begin{bmatrix}r_1\\s_1\end{bmatrix}=r_0\begin{bmatrix}0.95\\0.05\end{bmatrix}+s_0\begin{bmatrix}0.03\\0.97\end{bmatrix}=\begin{bmatrix}0.95&0.03\\0.05&0.97\end{bmatrix}\begin{bmatrix}r_0\\s_0\end{bmatrix}$$

即

$$\boldsymbol{x}_1=M\boldsymbol{x}_0\tag{8}$$

其中 M 为**迁移矩阵**(migration matrix)，它由下表确定：

$$\begin{matrix}&\text{从：}&\\\text{城市}&\text{郊区}&\text{到：}\\\begin{bmatrix}0.95&0.03\\0.05&0.97\end{bmatrix}&&\begin{matrix}\text{城市}\\\text{郊区}\end{matrix}\end{matrix}$$

方程(8)表示从2000 年到2001 年的人口变化情况. 如果迁移百分比保持不变，则从2001 年到2002 年的人口变化由下面公式给出

$$\boldsymbol{x}_2=M\boldsymbol{x}_1$$

对2002 年到2003 年，以及往后年份也类似. 即有

$$\boldsymbol{x}_{k+1}=M\boldsymbol{x}_k\quad k=0,1,2,\cdots\tag{9}$$

向量序列$\{\boldsymbol{x}_0,\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots\}$描述了城市/郊区人口在若干年内的分布变化.

98

【例题 3】 已知2000 年的城市人口为600 000，郊区人口为400 000. 计算2001 年和2002 年的整个地区的人口.

解： 2000 年的初始人口为 $\boldsymbol{x}_0=\begin{bmatrix}600\ 000\\400\ 000\end{bmatrix}$. 对2001 年有

$$\boldsymbol{x}_1=\begin{bmatrix}0.95&0.03\\0.05&0.97\end{bmatrix}\begin{bmatrix}600\ 000\\400\ 000\end{bmatrix}=\begin{bmatrix}582\ 000\\418\ 000\end{bmatrix}$$

对2002 年有

$$\boldsymbol{x}_2=M\boldsymbol{x}_1=\begin{bmatrix}0.95&0.03\\0.05&0.97\end{bmatrix}\begin{bmatrix}582\ 000\\418\ 000\end{bmatrix}=\begin{bmatrix}565\ 440\\434\ 560\end{bmatrix}$$ ■

因为$\boldsymbol{x}_k\mapsto\boldsymbol{x}_{k+1}$是线性变换，(9)中人口迁移的模型是线性的. 线性依据两个事实，如(6)，(7)中所示，即选择从一个地区迁移到另一个地区的人口数量与这个地区的人口成正比，并且这些迁移累积的结果可以由从不同地区的人口迁移累加得到.

基础练习

求出矩阵 $\boldsymbol{A}$ 及向量 $\boldsymbol{x}$ 和 $\boldsymbol{b}$，使得例题 1 中的问题相当于求解 $\boldsymbol{Ax}=\boldsymbol{b}$.

---

1. 简单起见，我们忽略其他因素对人口的影响，例如人口的出生、死亡，以及城市/郊区与外界之间的迁入和迁出.

习题　1.10

1. 早餐的麦片包装盒上通常会列出一份麦片中所含卡路里，以及蛋白质、碳水化合物、脂肪的量．两种常见的麦片食品所含的卡路里及各种营养如下给出．

营　养	每份早餐所含的营养	
	General Mills 公司的 Cheerios 谷物早餐	Quaker 公司的 100% 纯天然谷物早餐
卡路里	110	130
蛋白质(g)	4	3
碳水化合物(g)	20	18
脂肪(g)	2	5

假设需要准备一份两种麦片的混合物，使之恰好包含 295 卡路里，9g 蛋白质，48g 碳水化合物和 8g 脂肪．

a. 为该问题建立一个向量方程．并说明方程中变量都表示什么．

99

b. 写出等价的矩阵方程，然后判断是否能得到所需的两种麦片的混合物．

2. 一份(28g) Kellogg's 脆燕麦片能提供 110 卡路里、3g 蛋白质、21 克碳水化合物以及 3g 脂肪．一份(28g) Kellogg's 脆片能提供 110 卡路里、2g 蛋白质、25 克碳水化合物以及 4g 脂肪．

a. 构造矩阵 B 和向量 u，使得 Bu 给出三份脆燕麦片与两份脆片的混合物所包含的卡路里、蛋白质、碳水化合物以及脂肪的总量．

b. [M]假设你需要一种谷物：蛋白质含量比脆片多，脂肪含量比脆燕麦片少．是否能得到提供 110 卡路里、2.25g 蛋白质、24g 碳水化合物和 1g 脂肪的两种谷物的混合物？如果可以，怎么混合？

3. 剑桥食谱除了表 1-6 中列出的营养之外，还提供了 0.8g 的钙．在剑桥食谱中的三种食品当中：一单位(100g)脱脂牛奶提供 1.26g 钙，一单位大豆粉提供 0.19g 钙，一单位乳清提供 0.8g 钙．在这份混合食谱中的另外一种食物是离析大豆蛋白，一个单位的离析大豆蛋白提供的营养如下：80g 蛋白质、0g 碳水化合物、3.4g 脂肪以及 0.18g 钙．

a. 建立一个矩阵方程，使该方程的解确定了脱脂牛奶、大豆粉、乳清以及离析大豆蛋白的总量，它们必须能提供剑桥食谱中蛋白质、碳水化合物、脂肪和钙的精确总量．说明方程中的变量都表示什么．

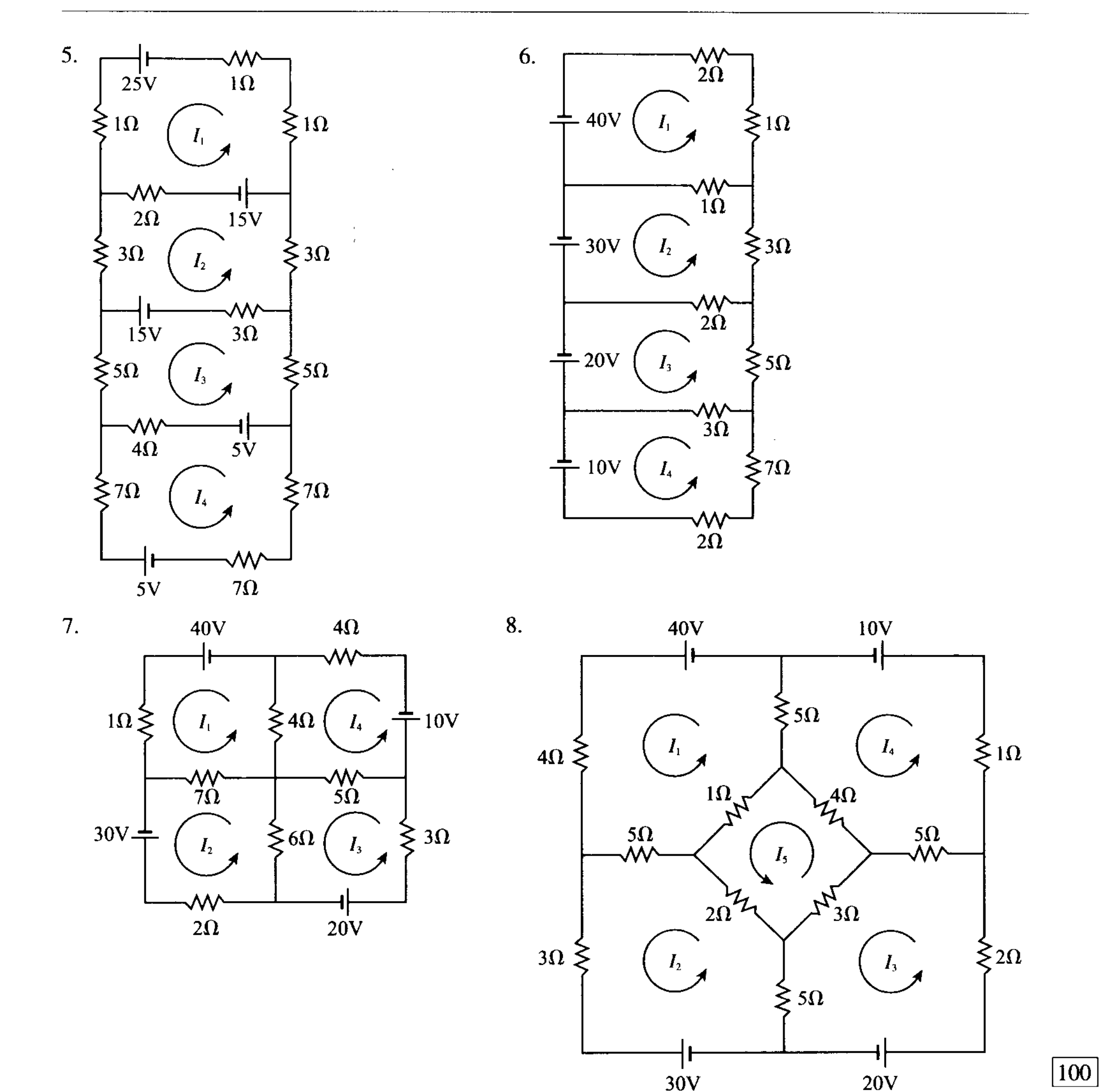
b. [M]求解(a)中的方程组，并讨论得到的答案．

4. 一个饮食学专家计划一份膳食，提供一定量的维生素 C、钙和镁．其中用到 3 种食物，它们的质量用适当的单位计量．这些食品提供的营养以及食谱需要的营养如下给出．

营　养	单位食物所含的营养(毫克)			需要的营养总量(毫克)
	食物 1	食物 2	食物 3	
维生素 C	10	20	20	100
钙	50	40	10	300
镁	30	10	40	200

针对这个问题写出一个向量方程．说明方程中的变量都表示什么，然后求解这个方程．

在习题 5 ~ 8 中，写出已知回路电流的矩阵方程．[M]如果有 MATLAB 或者其他的矩阵软件可以使用，那么求解这个回路电流的方程组．



100

9. 在某一个地区，每年大约有5%的城市人口移居到周围的郊区，大约4%的郊区人口移居到城市中．在2000年，城市中有600 000居民，郊区有400 000居民．建立一个差分方程来描述这种情况，用x_0表示2000年的初始人口．然后估计两年之后，即2002年城市和郊区的人口数量．（忽略其他因素对人口规模的影响．）

10. 在某一个地区，每年大约有7%的城市人口移居到周围的郊区，大约3%的郊区人口移居到城市中．在2000年，城市中有800 000居民，郊区有500 000居民．建立一个差分方程来描述这种情况，用x_0表示2000年的初始人口．然后估计两年之后，即2002年城市和郊区的人口数量．

11. 在1990年初，加利福尼亚的人口为29 716 000；而除加州以外的美国人口为218 994 000．在这一年间，有509 500人从加利福尼亚移居到美国的其他地方，同时有564 100人从美国

的其他地区移居到加利福尼亚.[1]

a. 根据这种情况，建立一个迁移矩阵，用 5 位小数表示加利福尼亚迁入和迁出的人口率. 说明是如何生成这个迁移矩阵的.

b. [M]通过计算，估计 2000 年加利福尼亚以及美国其他地区的人口数量，假设迁移率在 10 年时间中都没有变化(这个计算不考虑出生、死亡,以及从美国以外的地区迁入加利福尼亚或美国其他州的实际人数.).

12. [M]堪萨斯州 Wichita 市的 Budget Rent A Car 公司有一个车队，大约有 450 辆车，分布在三个地点. 一个地点租出去的车可以归还到任意三个地点. 下面矩阵给出汽车归还到每个地点的不同比率. 假设星期一在机场有 304 辆车(或从机场租出)，东部办公区有 48 辆车，西部办公区有 98 辆车. 那么在星期三时，车辆的大致分布是怎样?

车辆出租地			
机场	东部	西部	归还到
0.97	0.05	0.10	机场
0.00	0.90	0.05	东部
0.03	0.05	0.85	西部

13. [M]设 $\boldsymbol{M}$ 和 $\boldsymbol{x}_0$ 同例题 3.

a. 计算人口向量 $\boldsymbol{x}_k, k=1,\cdots,20$. 你发现了什么?

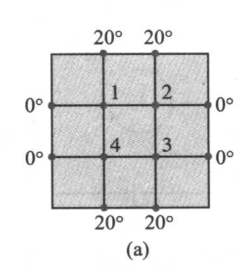
b. 当城市人口为 350 000，郊区人口为 650 000 时，重做(a). 你发现了什么?

14. [M]研究金属薄板边界温度的变化如何影响薄板内部点的温度.

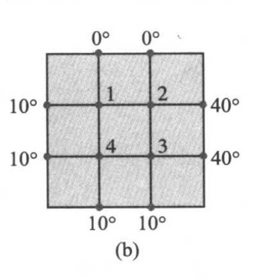
a. 先估计图中所示的金属薄板上 4 点的温度 T_1，…，T_4. 在每一种情况中，T_k的值近似等于周围四个最近节点的平均温度. 1.1 节习题 33 和 34 中得到的值(度)为(20,27.5,30, 22.5). 这些值与你从(a)和(b)中得到的结果有什么关系?

b. 不做任何计算，猜想当边界温度都乘上 3 时，(a)中内部点的温度. 验证你的猜测.

c. 最后，对于从八个边界点温度到四个内部点温度的对应关系，给出一个一般推测.



(a)



(b)

101

基础练习答案

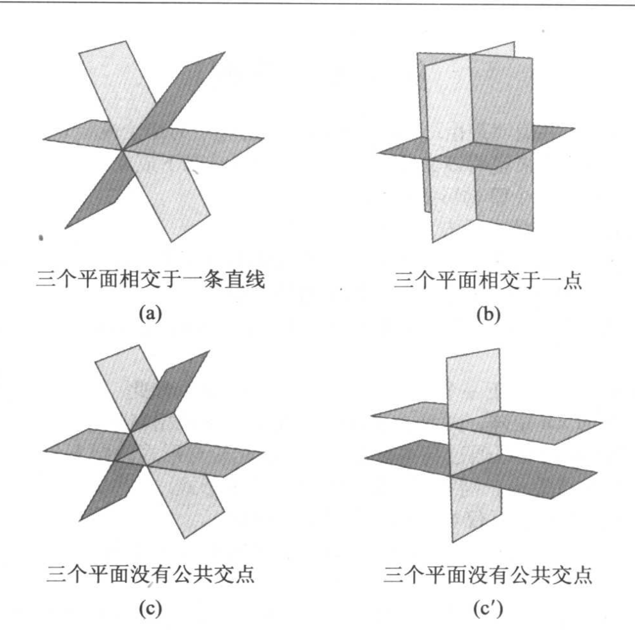
$$\boldsymbol{A}=\begin{bmatrix}36 & 51 & 13\\ 52 & 34 & 74\\ 0 & 7 & 1.1\end{bmatrix},\quad \boldsymbol{x}=\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix},\quad \boldsymbol{b}=\begin{bmatrix}33\\ 45\\ 3\end{bmatrix}$$

1. 移民数据由加州金融财政部人口统计研究联合会提供.

第 1 章补充题

1. 判断各命题的真假，并说明理由.（如果正确，指出应用的结论或定理. 如果错误，解释为什么，或给出一个反例来说明该命题为什么不在所有情况下成立.）
   a. 每个矩阵都与唯一一个阶梯形式矩阵行等价.
   b. 含 n 个变量 n 个方程的线性方程组至多有 n 个解.
   c. 如果一个线性方程组有两个不同的解，那么它必然有无穷多解.
   d. 如果一个线性方程组没有自由变量，则它有唯一解.
   e. 如果增广矩阵 $[A\ b]$ 通过基本行变换变为 $[C\ b]$，则方程 $Ax=b$ 和 $Cx=b$ 有相同的解集.
   f. 如果方程组 $Ax=b$ 有多于一个的解，则方程组 $Ax=0$ 也如此.
   g. 如果 A 为一个 $m\times n$ 矩阵，并且方程 $Ax=b$ 对某些 b 是相容的. 则 A 中的列张成 $\mathbf{R}^n$.
   h. 如果增广矩阵 $[A\ b]$ 可以通过基本行变换变为简化阶梯形式，则方程 $Ax=b$ 是相容的.
   i. 如果矩阵 A 和 B 行等价，则它们有相同的简化阶梯形式.
   j. 方程 $Ax=0$ 有平凡解当且仅当方程中没有自由变量.
   k. 如果 A 为一个 $m\times n$ 矩阵，并且方程 $Ax=b$ 对所有的 b 都相容. 则 A 中含有 m 个主元列.
   l. 如果一个 $m\times n$ 矩阵 A 中的每一行都有主元位置，则方程 $Ax=b$ 对 $\mathbf{R}^m$ 中的任意 b 都只有唯一解.
   m. 如果一个 $n\times n$ 矩阵 A 中含有 n 个主元位置，则 A 的简化阶梯形式为一个 $n\times n$ 单位矩阵.
   n. 如果 3×3 矩阵 A 和 B 中都含有 3 个主元位置，则 A 可以通过初等行变换变为 B.
   o. A 为一个 $m\times n$ 矩阵，并且方程 $Ax=b$ 至少有两个不同的解. 如果方程 $Ax=c$ 相容，则方程 $Ax=c$ 有多个解.
   p. 如果矩阵 A 和 B 是行等价的 $m\times n$ 矩阵，且 A 中的列张成 $\mathbf{R}^m$. 则 B 中的列也张成 $\mathbf{R}^m$.
   q. 如果 $\mathbf{R}^3$ 中向量集合 $S=\{v_1,v_2,v_3\}$ 中没有互为倍数的向量，则 S 是线性无关集.
   r. 如果 $\{u,v,w\}$ 线性无关，则 u，v，w 不在 $\mathbf{R}^2$ 中.
   s. 在某些情况下，四个向量有可能张成 $\mathbf{R}^5$.
   t. 如果 u，v 为 $\mathbf{R}^m$ 中向量，则 $-u$ 在 $\mathrm{Span}\{u,v\}$ 中.
   u. 设 u，v，w 为 $\mathbf{R}^2$ 中非零向量，则 w 为 u，v 的线性组合.
   v. 如果 w 为 $\mathbf{R}^n$ 中 u，v 的线性组合，则 u 为 v，w 的线性组合.
   w. 假设 v_1，v_2，v_3 为 $\mathbf{R}^5$ 中向量，v_2 不是 v_1 的倍数，v_3 不是 v_1 和 v_2 的线性组合. 则 $\{v_1,v_2,v_3\}$ 是线性无关集.
   x. 一个线性变换是一个函数.
   y. 如果 A 是 6×5 矩阵，线性变换 $x\mapsto Ax$ 不能把 $\mathbf{R}^5$ 映射到 $\mathbf{R}^6$ 上.
   z. 如果 A 是含有 m 个主元位置的 $m\times n$ 矩阵，则线性变换 $x\mapsto Ax$ 是一对一映射.
2. 设 a 和 b 为两个实数. 写出(线性)方程 $ax=b$ 可能的解集. [提示：解的个数依赖于 a 和 b.]
3. 当 a，b，c 不全为零时，单个线性方程 $ax+by+cz=d$ 的解 (x,y,z) 构成中 $\mathbf{R}^3$ 的一个平面. 构造一个含有三个线性方程的集合，使它们的图像(a)相交于同一条直线，(b)相交于一点，(c)没有交点. 典型的图像如图 1-48 所示.

102



三个平面相交于一条直线 (a)

三个平面相交于一点 (b)

三个平面没有公共交点 (c)

三个平面没有公共交点 (c′)

图 1-48

4. 如果一个含有 3 个未知量三个方程的线性方程组，其系数矩阵的每一列中都有一个主元位置. 解释为什么该方程组有唯一解.
5. 确定 h 和 k 的取值，使方程组的解集(i)为空，(ii)有唯一解，(iii)有无穷多解.

   a. $\begin{aligned} x_1+3x_2&=k \\ 4x_1+hx_2&=8 \end{aligned}$　　b. $\begin{aligned} -2x_1+hx_2&=1 \\ 6x_1+kx_2&=-2 \end{aligned}$
6. 考虑如下的方程组是否相容：

$$\begin{aligned} 4x_1-2x_2+7x_3&=-5 \\ 8x_1-3x_2+10x_3&=-3 \end{aligned}$$

   (a)定义适当的向量，用线性组合的形式重述这个问题，并解决.

   (b)定义适当的矩阵 $\boldsymbol{A}$，用“$\boldsymbol{A}$ 中的列”来重述这个问题.

   (c)用(b)中的矩阵定义适当的线性变换 T，然后用变换 T 的形式重述这个问题.
7. 考虑如下方程组是否对所有的 b_1，b_2，b_3都相容：

$$\begin{aligned} 2x_1-4x_2-2x_3&=b_1 \\ -5x_1+x_2+x_3&=b_2 \\ 7x_1-5x_2-3x_3&=b_3 \end{aligned}$$

   (a)定义适当的向量 $\boldsymbol{v}_1$，$\boldsymbol{v}_2$，$\boldsymbol{v}_3$，用 $\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$的形式重述这个问题，并解决.

   (b)定义适当的矩阵 $\boldsymbol{A}$，用“$\boldsymbol{A}$ 中的列.”来重述这个问题.

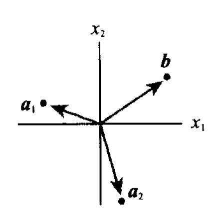
   (c)用(b)中的矩阵定义适当的线性变换 T，然后用变换 T 的形式重述这个问题.
8. 利用 1.2 节中习题 1 的符号，表示矩阵 $\boldsymbol{A}$ 可能的阶梯形式.

   a. $\boldsymbol{A}$ 是 2×3 矩阵，且 $\boldsymbol{A}$ 中的列张成 $\mathbf{R}^2$.

   b. $\boldsymbol{A}$ 是 3×3 矩阵，且 $\boldsymbol{A}$ 中的列张成 $\mathbf{R}^3$.
9. 把向量 $\begin{bmatrix}5\\6\end{bmatrix}$写成两个向量和的形式，一个向量在直线 $\{(x,y):y=2x\}$上，另一个在直线

$\{(x,y):y=x/2\}$上.

10. 设 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{b}$ 为 $\mathbf{R}^2$ 中向量,如右图所示. 且设 $A=[\boldsymbol{a}_1\quad \boldsymbol{a}_2]$. 方程 $A\boldsymbol{x}=\boldsymbol{b}$ 是否有解? 如果有,解是否唯一? 请解释.



11. 构造一个不为阶梯形式的 2×3 矩阵 A,使 $A\boldsymbol{x}=\boldsymbol{0}$ 的解为 $\mathbf{R}^3$ 中的一条直线.
12. 构造一个不为阶梯形式的 2×3 矩阵 A,使 $A\boldsymbol{x}=\boldsymbol{0}$ 的解为 $\mathbf{R}^3$ 中的一个平面.
13. 写出一个 3×3 矩阵 A 的简化阶梯形式,使 A 的前两列为主元列,并且

$$A\begin{bmatrix}3\\-2\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

14. 确定 a 的取值,使$\left\{\begin{bmatrix}1\\a\end{bmatrix},\begin{bmatrix}a\\a+2\end{bmatrix}\right\}$线性无关.
15. 在(a)和(b)中,假设向量组都是线性无关的. 对于 $a,\cdots,f$ 你能得到什么结论? 给出解释. [提示:(b)需要用到一个定理.]

(a) $\begin{bmatrix}a\\0\\0\end{bmatrix},\begin{bmatrix}b\\c\\0\end{bmatrix},\begin{bmatrix}d\\e\\f\end{bmatrix}$　　(b) $\begin{bmatrix}a\\1\\0\\0\end{bmatrix},\begin{bmatrix}b\\c\\1\\0\end{bmatrix},\begin{bmatrix}d\\e\\f\\1\end{bmatrix}$

16. 利用 1.7 中的定理 7 解释 A 中的列为什么线性无关.

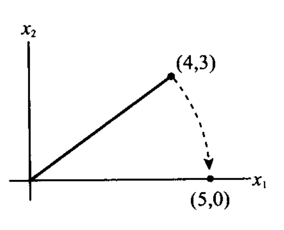
$$A=\begin{bmatrix}1&0&0&0\\2&5&0&0\\3&6&8&0\\4&7&9&10\end{bmatrix}$$

17. 解释为什么当$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$线性无关,且 $\boldsymbol{v}_4$ 不在 Span$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$中时,$\mathbf{R}^5$ 中的向量集合$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4\}$一定是线性相关的.

103

18. 假设$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$是 $\mathbf{R}^n$ 中的线性无关集合. 说明$\{\boldsymbol{v}_1,\boldsymbol{v}_1+\boldsymbol{v}_2\}$也是线性无关的.
19. 假设 $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$ 为 $\mathbf{R}^3$ 中一条直线上不同的点. 这条直线不一定要经过原点. 证明$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$线性相关.
20. 设 T: $\mathbf{R}^n\to\mathbf{R}^m$ 为一个线性变换,且 $T(\boldsymbol{u})=\boldsymbol{v}$. 证明 $T(-\boldsymbol{u})=-\boldsymbol{v}$.
21. 设线性变换 T: $\mathbf{R}^3\to\mathbf{R}^3$ 将向量做关于平面 $x_2=0$ 的反射,即 $T(x_1,x_2,x_3)=(x_1,-x_2,x_3)$. 求出变换 T 的标准矩阵.
22. 设 A 为 3×3 矩阵,且线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 把 $\mathbf{R}^3$ 映射到 $\mathbf{R}^3$ 上. 为什么这个线性变换一定是一对一的?
23. 计算机程序中的吉文斯旋转是一个从 $\mathbf{R}^n$ 到 $\mathbf{R}^n$ 上线性变换,该旋转生成向量(通常是矩阵中的一列)中的一个零元素. $\mathbf{R}^2$ 中吉文斯旋转的标准矩阵有如下形式

$$\begin{bmatrix}a&-b\\b&a\end{bmatrix},\ a^2+b^2=1$$



习题23 图　$\mathbf{R}^2$ 中的一个吉文斯旋转

求出 a 和 b，使得 $\begin{bmatrix}4\\3\end{bmatrix}$ 旋转得到 $\begin{bmatrix}5\\0\end{bmatrix}$.

24. 下面的方程组表示了 $\mathbf{R}^2$ 中的一个吉文斯旋转．求出 a 和 b.

$$\begin{bmatrix}a & 0 & -b\\0 & 1 & 0\\b & 0 & a\end{bmatrix}\begin{bmatrix}2\\3\\4\end{bmatrix}=\begin{bmatrix}2\sqrt{5}\\3\\0\end{bmatrix},\quad a^2+b^2=1$$

25. 建造一个大的公寓楼将用到模块构造技术．每一个楼层的房间安排将从 3 个基本的楼层计划中选择．计划 A 的一层楼中有 18 套房，包括 3 个三居，7 个两居，以及 8 个一居．计划 B 的一层楼中有 4 个三室单元，4 个两室单元，以及 8 个一室单元．计划 C 的一层楼中有 5 个三室单元，3 个两室单元，以及 9 个一室单元．假设该建筑总共有 x_1 层应用计划 A，x_2 层应用计划 B，x_3 层应用计划 C.

a. 对向量 $x_1\begin{bmatrix}3\\7\\8\end{bmatrix}$，你能给出什么解释？

b. 写出一个正式的向量线性组合，这些向量表示出建筑中包含的三室、两室、一室套房的总数.

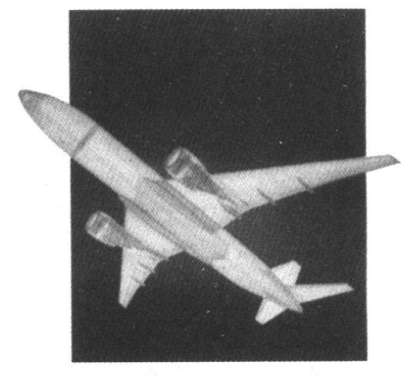
c. [**M**]可以设计出含有 66 个三室单元、74 个两室单元、136 个一室单元的建筑吗？如果可以，是否不止一种方法可以做到？请解释你的答案.

104

第2章 矩阵代数

实例介绍：飞机设计中的计算机模型

为了设计下一代民用和军用飞机，波音虚拟工作室的工程师们使用了3D建模方法和计算流体力学(CFD). 在建造飞机的物理模型之前，他们研究虚拟飞机周围的气流以解决设计中的重大疑问. 这在很大程度上缩减了设计周期和费用，其中线性代数发挥了至关重要的作用.



虚拟飞机开始只是一个存储在计算机内，呈现在图形显示终端上的“曲线框架”数学模型.(本页给出了波音777的一个模型.)这个数学模型组织并影响飞机(不论是外部还是内部)设计和制造的每一步. CFD分析则主要考虑飞机外部的设计和制造.

虽然最后完成的飞机表面看似平滑，其几何结构却错综复杂. 除了机翼和机身，一架飞机还包括发动机舱、水平尾翼、活动扶助翼和副翼. 空气在这些部件周围的流动方式决定了飞机在天空中的飞行方式. 描述这些气流的方程很复杂，其中还必须考虑发动机的吸气量和排气量，以及机翼留下的尾迹. 因此，为了研究气流，工程师需要高度精确地描述飞机的表面.

为了得到飞机表面的模型，计算机首先向原有的“曲线框架”模型添加一个三维的立方体网格. 网格中的立方体或者完全位于飞机内部，或者完全位于飞机外部，或者与飞机表面相交. 计算机选择那些与飞机表面相交的立方体并且细分它们，然后保留仍与飞机表面相交的那些小立方体. 重复这一细分过程，直到网格变得足够精细. 一个典型的网格可能包含逾400 000个立方体.

计算飞机表面的气流需要反复求解包含多达两百万个方程和变量的线性方程组 $A\mathbf{x}=\mathbf{b}$. 向量 $\mathbf{b}$ 每次都随网格数据以及前面方程的解而变化. 使用市场上速度最快的计算机，波音虚拟工作室的一个小组建立并求解一个气流问题就要花费几小时到几天的时间. 对解进行分析以后，他们可能需要对飞机表面进行小的修改，然后重复上面整个计算过程，累计起来共需要处理数千个CFD问题.

本章介绍的下面两个重要概念将有助于求解这类大型方程组：

- 分块矩阵：一个典型的CFD方程组具有“稀疏”的系数矩阵，即其中的元素大多数都为零. 对变量进行适当分组，就能得到包含很多零块的分块矩阵. 2.4节将介绍这类矩阵并且给出它们的一些应用.
- 矩阵分解：即使写成了分块矩阵，这类方程组仍然十分复杂. 为了进一步化简计算，波音虚拟工作室在其所用的CFD软件中运用了系数矩阵的LU分解

技术. 2.5 节讨论了 LU 分解和其他一些有用的矩阵分解. 有关矩阵分解的更多内容将在后面几章中给出.

为了分析气流方程组的解，工程师们试图模拟飞机表面的气流. 他们使用了以线性代数为实现机制的计算机绘图软件. 飞机表面的“曲线框架”模型以数据矩阵的形式存储在计算机中. 当图像在计算机屏幕上呈现时，工程师们可以改变显示比例、缩小或放大局部，或者旋转图像以便看到隐藏于视野之外的部分. 每一种操作都由相应的矩阵乘法来完成，其基本思想将在 2.7 节予以解释.

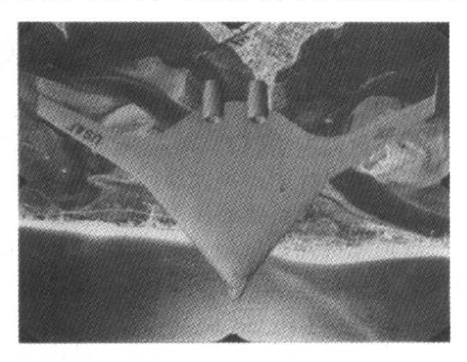


图 2-1 现代 CFD 革新了机翼的设计. 波音混合机翼体正在设计中，最迟于 2020 年完成

一旦掌握了矩阵的代数运算，我们分析和解决方程组的能力将会大大增强. 本章所介绍的定义和定理还将为我们提供一些基本工具，使我们能够处理涉及两个或多个矩阵的线性代数应用问题. 对于方阵，2.3 节 106 的可逆矩阵定理与本书前面介绍的大部分概念相呼应. 2.4 节和 2.5 节考察了线性代数中应用十分广泛的分块矩阵和矩阵分解. 2.6 节和 2.7 节叙述了矩阵代数在经济和计算机绘图方面的两个有趣的应用.

2.1 矩阵运算

如果 $\boldsymbol{A}$ 是一个 $m\times n$ 矩阵，即 m 行 n 列矩阵，那么 $\boldsymbol{A}$ 的第 i 行第 j 列的元素记为 a_{ij}，称为 $\boldsymbol{A}$ 的 (i,j) – 元. 见图 2-2. 例如，$(3,2)$ – 元是第 3 行第 2 列上的 a_{32}. $\boldsymbol{A}$ 的列是 $\mathbf{R}^m$ 中的向量，记为 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$. 当我们把 $\boldsymbol{A}$ 写成

$$\boldsymbol{A}=[\boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \cdots\ \ \boldsymbol{a}_n]$$

时，我们就只关注这些列. 注意，a_{ij} 是列向量 $\boldsymbol{a}_j$ 的第 i 个元素(从上往下数).

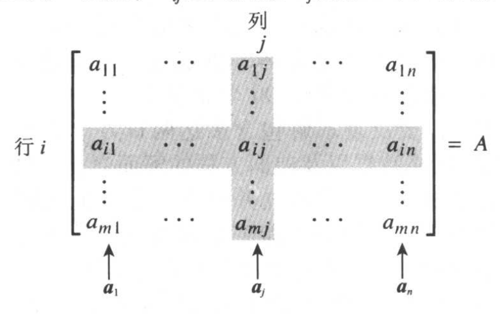


图 2-2 矩阵记号

$m\times n$ 矩阵 $\boldsymbol{A}=[a_{ij}]$ 的**对角线元素**(diagonal entries)是 $a_{11},a_{22},a_{33},\cdots$，它们构成 $\boldsymbol{A}$ 的**主对角线**(main diagonal). **对角矩阵**(diagonal matrix)是非对角线上元素全为零的方阵. $n\times n$ 单位矩阵 $\boldsymbol{I}_n$ 就是一个对角矩阵. 元素全为零的 $m\times n$ 矩阵是**零矩阵**(zero

matrix)，记作 $\mathbf{0}$. 零矩阵的维度可以通过上下文看出.

2.1.1 矩阵的和与数量倍数

先前介绍的向量算术可以自然地推广到矩阵. 我们称两个矩阵**相等**(equal)，如果它们维度相同(即有相同的行数和列数)且对应列相等，即它们的对应元素相等. 如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是 $m\times n$ 矩阵，则**和**(sum) $\boldsymbol{A}+\boldsymbol{B}$ 也是 $m\times n$ 矩阵，它的列是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 中对应列之和. 列向量加法是逐个元素相加，所以 $\boldsymbol{A}+\boldsymbol{B}$ 的每一个元素都是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 中对应元素之和. 仅当 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的维度相同时，和 $\boldsymbol{A}+\boldsymbol{B}$ 才有定义.

107

【例题 1】 设

$$A=\begin{bmatrix}4&0&5\\-1&3&2\end{bmatrix},\quad B=\begin{bmatrix}1&1&1\\3&5&7\end{bmatrix},\quad C=\begin{bmatrix}2&-3\\0&1\end{bmatrix}$$

则

$$A+B=\begin{bmatrix}5&1&6\\2&8&9\end{bmatrix}$$

但是 A + C 没有定义，因为 A 和 C 的维度不同. ■

如果 r 是数量，$\boldsymbol{A}$ 是矩阵，那么**数乘**(scalar multiple) $r\boldsymbol{A}$ 是矩阵，它的列是 $\boldsymbol{A}$ 中对应列的 r 倍. 与向量记法相同，我们定义 $(-1)\boldsymbol{A}$ 为 $-\boldsymbol{A}$，并且以 $\boldsymbol{A}-\boldsymbol{B}$ 记 $\boldsymbol{A}+(-1)\boldsymbol{B}$.

【例题 2】 如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是例题 1 中的矩阵，则

$$2\boldsymbol{B}=2\begin{bmatrix}1&1&1\\3&5&7\end{bmatrix}=\begin{bmatrix}2&2&2\\6&10&14\end{bmatrix}$$

$$\boldsymbol{A}-2\boldsymbol{B}=\begin{bmatrix}4&0&5\\-1&3&2\end{bmatrix}-\begin{bmatrix}2&2&2\\6&10&14\end{bmatrix}=\begin{bmatrix}2&-2&3\\-7&-7&-12\end{bmatrix}$$ ■

例题 2 中没必要把 $\boldsymbol{A}-2\boldsymbol{B}$ 按 $\boldsymbol{A}+(-2)\boldsymbol{B}$ 来计算，因为通常的代数运算法则也适用于矩阵的加法和数乘，如下面定理所述.

【定理 1】 设 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 是维度相同的矩阵，r 和 s 是数量.

a. $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$

b. $(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}=\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$

c. $\boldsymbol{A}+\boldsymbol{0}=\boldsymbol{A}$

d. $r(\boldsymbol{A}+\boldsymbol{B})=r\boldsymbol{A}+r\boldsymbol{B}$

e. $(r+s)\boldsymbol{A}=r\boldsymbol{A}+s\boldsymbol{A}$

f. $r(s\boldsymbol{A})=(rs)\boldsymbol{A}$

定理 1 中的每一个等式都可通过验证左端矩阵和右端矩阵维度相同、对应列相等来得到. $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 维度相同已经知道，对应列相等可以由向量的性质类推得到. 例如，如果 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 的第 j 列分别是 $\boldsymbol{a}_j$，$\boldsymbol{b}_j$，$\boldsymbol{c}_j$，则 $(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}$ 和 $\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$ 的第 j 列分别是

$$(\boldsymbol{a}_j+\boldsymbol{b}_j)+\boldsymbol{c}_j \qquad \text{和} \qquad \boldsymbol{a}_j+(\boldsymbol{b}_j+\boldsymbol{c}_j)$$

因为对每个 j 这两个向量和都相等，性质(b)得证.

根据加法结合律，我们可以把 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 的和简写为 $\boldsymbol{A}+\boldsymbol{B}+\boldsymbol{C}$，它既可以按 $(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}$ 计算，也可以按 $\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$ 计算. 这个记法对 4 个或更多矩阵的加法也适用.

108

2.1.2 矩阵乘法

用矩阵 $\boldsymbol{B}$ 乘向量 $\boldsymbol{x}$ 时，$\boldsymbol{B}$ 将 $\boldsymbol{x}$ 变换到向量 $\boldsymbol{B}\boldsymbol{x}$. 如果这个向量再乘以矩阵 $\boldsymbol{A}$，最

终得到的矩阵为 $\boldsymbol{A}(\boldsymbol{Bx})$. 如图2-3.

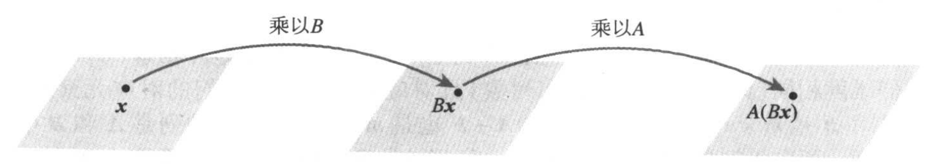


图2-3　依次乘以 $\boldsymbol{B}$ 和 $\boldsymbol{A}$

因此，$\boldsymbol{A}(\boldsymbol{Bx})$ 是由 $\boldsymbol{x}$ 经映射(即1.8节所讨论的线性变换)的复合得到，我们的目标是把这个复合映射表示成单个矩阵的乘法，以 $\boldsymbol{AB}$ 记这个矩阵，则有

$$\boldsymbol{A}(\boldsymbol{Bx}) = (\boldsymbol{AB})\boldsymbol{x} \tag{1}$$

如图2-4.

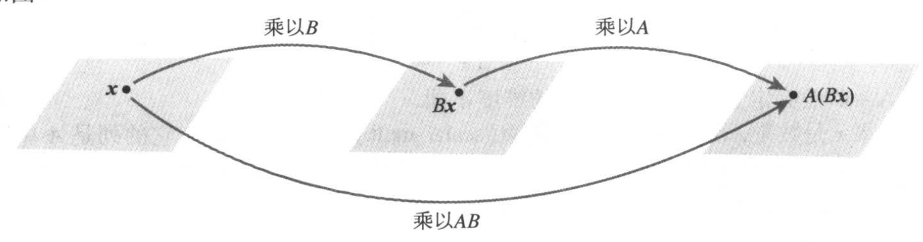


图2-4　乘以 $\boldsymbol{AB}$

如果 $\boldsymbol{A}$ 是 $m \times n$ 矩阵，$\boldsymbol{B}$ 是 $n \times p$ 矩阵且 $\boldsymbol{x} \in \mathbf{R}^p$，记 $\boldsymbol{B}$ 的列向量为 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p$，向量 $\boldsymbol{x}$ 的元素为 $x_1,\cdots,x_p$. 则

$$\boldsymbol{Bx} = x_1\boldsymbol{b}_1 + \cdots + x_p\boldsymbol{b}_p$$

由于乘以 $\boldsymbol{A}$ 的运算是线性的，所以

$$\begin{aligned}\boldsymbol{A}(\boldsymbol{Bx}) &= \boldsymbol{A}(x_1\boldsymbol{b}_1) + \cdots + \boldsymbol{A}(x_p\boldsymbol{b}_p)\\ &= x_1\boldsymbol{Ab}_1 + \cdots + x_p\boldsymbol{Ab}_p\end{aligned}$$

即向量 $\boldsymbol{A}(\boldsymbol{Bx})$ 是向量 $\boldsymbol{Ab}_1,\cdots,\boldsymbol{Ab}_p$ 的一个线性组合，其权重为 $\boldsymbol{x}$ 中的元素. 如果我们把这些向量重新写成一个矩阵的列向量，则有

$$\boldsymbol{A}(\boldsymbol{Bx}) = [\boldsymbol{Ab}_1 \quad \boldsymbol{Ab}_2 \quad \cdots \quad \boldsymbol{Ab}_p]\boldsymbol{x}$$

109　乘以 $[\boldsymbol{Ab}_1 \quad \boldsymbol{Ab}_2 \quad \cdots \quad \boldsymbol{Ab}_p]$ 的运算将 $\boldsymbol{x}$ 变换到 $\boldsymbol{A}(\boldsymbol{Bx})$. 它就是我们要求的矩阵！

【定义】　如果 $\boldsymbol{A}$ 是 $m \times n$ 矩阵，$\boldsymbol{B}$ 是列向量为 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p$ 的 $n \times p$ 矩阵，则乘积 $\boldsymbol{AB}$ 是 $m \times p$ 矩阵，其列向量为 $\boldsymbol{Ab}_1,\cdots,\boldsymbol{Ab}_p$. 即，

$$\boldsymbol{AB} = \boldsymbol{A}[\boldsymbol{b}_1 \quad \boldsymbol{b}_2 \quad \cdots \quad \boldsymbol{b}_p] = [\boldsymbol{Ab}_1 \quad \boldsymbol{Ab}_2 \quad \cdots \quad \boldsymbol{Ab}_p]$$

这个定义使得(1)对任意 $\boldsymbol{x} \in \mathbf{R}^p$ 都成立. 等式(1)说明图2-4中的复合映射是线性变换并且其标准矩阵是 $\boldsymbol{AB}$. 矩阵的乘法对应线性变换的复合.

【例题3】　计算 $\boldsymbol{AB}$，其中 $\boldsymbol{A} = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ 且 $\boldsymbol{B} = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$.

解： 记 $\boldsymbol{B} = [\boldsymbol{b}_1 \ \ \boldsymbol{b}_2 \ \ \boldsymbol{b}_3]$，计算

$$A\boldsymbol{b}_1=\begin{bmatrix}2&3\\1&-5\end{bmatrix}\begin{bmatrix}4\\1\end{bmatrix}=\begin{bmatrix}11\\-1\end{bmatrix},\quad A\boldsymbol{b}_2=\begin{bmatrix}2&3\\1&-5\end{bmatrix}\begin{bmatrix}3\\-2\end{bmatrix}=\begin{bmatrix}0\\13\end{bmatrix},\quad A\boldsymbol{b}_3=\begin{bmatrix}2&3\\1&-5\end{bmatrix}\begin{bmatrix}6\\3\end{bmatrix}=\begin{bmatrix}21\\-9\end{bmatrix}$$

则

$$AB=A[\boldsymbol{b}_1\quad \boldsymbol{b}_2\quad \boldsymbol{b}_3]=\begin{bmatrix}11&0&21\\-1&13&-9\end{bmatrix}$$

$$\uparrow\ A\boldsymbol{b}_1\quad \uparrow\ A\boldsymbol{b}_2\quad \uparrow\ A\boldsymbol{b}_3$$

■

注意到 $\boldsymbol{AB}$ 的第一列是 $A\boldsymbol{b}_1$，它是 $\boldsymbol{A}$ 中列的一个线性组合，其权重为 $\boldsymbol{b}_1$ 的元素. $\boldsymbol{AB}$ 的其他各列也类似.

$\boldsymbol{AB}$ 的每一列都是 $\boldsymbol{A}$ 中各列的一个线性组合，其权重为 $\boldsymbol{B}$ 中对应列上的元素.

显然，为了使线性组合 $A\boldsymbol{b}_1$ 有意义，$\boldsymbol{A}$ 的列数必须等于 $\boldsymbol{B}$ 的行数. 此外，$\boldsymbol{AB}$ 的定义说明 $\boldsymbol{AB}$ 与 $\boldsymbol{A}$ 有相同的行数，与 $\boldsymbol{B}$ 有相同的列数.

【例题 4】 $\boldsymbol{A}$ 是 3×5 矩阵，$\boldsymbol{B}$ 是 5×2 矩阵，如果 $\boldsymbol{AB}$ 和 $\boldsymbol{BA}$ 有意义，则它们的维度分别是多少？

110

解： 因为 $\boldsymbol{A}$ 有 5 列，$\boldsymbol{B}$ 有 5 行，所以乘积 $\boldsymbol{AB}$ 有意义，并且是 3×2 矩阵：

$$\underset{3\times5}{\overset{A}{\begin{bmatrix}*&*&*&*&*\\ *&*&*&*&*\\ *&*&*&*&*\end{bmatrix}}}\ \underset{5\times2}{\overset{B}{\begin{bmatrix}*&*\\ *&*\\ *&*\\ *&*\\ *&*\end{bmatrix}}}=\underset{3\times2}{\overset{AB}{\begin{bmatrix}*&*\\ *&*\\ *&*\end{bmatrix}}}$$

相等

$\boldsymbol{AB}$的维度

乘积 $\boldsymbol{BA}$ 没有定义，因为 $\boldsymbol{B}$ 的列数 2 不等于 $\boldsymbol{A}$ 的行数 3. ■

$\boldsymbol{AB}$ 的定义在理论和应用上都很重要，不过，当用手算解答小型问题时，下面的法则为计算 $\boldsymbol{AB}$ 中的单个元素提供了一种更有效的方法.

计算 $\boldsymbol{AB}$ 的行列法则

如果乘积 $\boldsymbol{AB}$ 有意义，则 $\boldsymbol{AB}$ 第 i 行、第 j 列上的元素是 $\boldsymbol{A}$ 的第 i 行与 $\boldsymbol{B}$ 的第 j 列上对应元素乘积之和. 若以 $(\boldsymbol{AB})_{ij}$ 记 $\boldsymbol{AB}$ 的 (i,j) – 元，且 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，则

$$(\boldsymbol{AB})_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj}$$

为验证这个法则，令 $\boldsymbol{B}=[\boldsymbol{b}_1\cdots\boldsymbol{b}_p]$. $\boldsymbol{AB}$ 的第 j 列是 $A\boldsymbol{b}_j$，我们可以用 1.4 节中计算 $A\boldsymbol{x}$ 的行向量法则来计算 $A\boldsymbol{b}_j$. $A\boldsymbol{b}_j$ 的第 i 个元素是 $\boldsymbol{A}$ 的第 i 行与向量 $\boldsymbol{b}_j$ 中对应元素的乘积之和，这正是上面法则所述的 $\boldsymbol{AB}$ 中 $(i,\ j)$ – 元的算法.

【例题 5】 用行列法则计算例题 3 中矩阵 $\boldsymbol{AB}$ 的两个元素. 仔细观察计算过程中数

值的变化，你将发现用两种方法计算 AB 都得到同一个矩阵.

解： 为求 $\boldsymbol{AB}$ 第 1 行、第 3 列上的元素，考虑 $\boldsymbol{A}$ 的第 1 行和 $\boldsymbol{B}$ 的第 3 列. 让对应元素相乘，然后累加乘积，如下所示：

111

$$\boldsymbol{AB} = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 2(6)+3(3) \\ \square & \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & \square & \square \end{bmatrix}$$

为求 $\boldsymbol{AB}$ 第 2 行、第 2 列上的元素，利用 $\boldsymbol{A}$ 的第 2 行和 $\boldsymbol{B}$ 的第 2 列：

$$\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 1(3)+-5(-2) & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 13 & \square \end{bmatrix} \quad \blacksquare$$

【例题 6】 计算 $\boldsymbol{AB}$ 第 2 行的元素，其中

$$\boldsymbol{A} = \begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix}$$

解： 由行列法则，$\boldsymbol{AB}$ 第 2 行的元素应该由 $\boldsymbol{A}$ 的第 2 行与 $\boldsymbol{B}$ 的各列计算得到：

$$\begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix}\begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} \square & \square \\ -4+21-12 & 6+3-8 \\ \square & \square \\ \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square \\ 5 & 1 \\ \square & \square \\ \square & \square \end{bmatrix} \quad \blacksquare$$

注意，例题 6 只需计算 $\boldsymbol{AB}$ 第 2 行的元素，我们可以在 $\boldsymbol{B}$ 的左边只写出 $\boldsymbol{A}$ 的第 2 行，计算

$$\begin{bmatrix} -1 & 3 & -4 \end{bmatrix}\begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 1 \end{bmatrix}$$

关于 $\boldsymbol{AB}$ 行的这个发现在一般情形下也成立，它可由行列法则推知. 以 $\text{row}_i(\boldsymbol{A})$ 记矩阵 $\boldsymbol{A}$ 的第 i 行. 则

$$\text{row}_i(\boldsymbol{AB}) = \text{row}_i(\boldsymbol{A}) \cdot \boldsymbol{B} \tag{2}$$

2.1.3 矩阵乘法的性质

下面的定理列举了矩阵乘法的标准性质. 回忆 $\boldsymbol{I}_m$ 表示 $m \times m$ 单位矩阵，且对任意 $\boldsymbol{x} \in \mathbf{R}^m$ 都有 $\boldsymbol{I}_m\boldsymbol{x} = \boldsymbol{x}$.

112

【定理 2】 设 $\boldsymbol{A}$ 是 $m \times n$ 矩阵，且 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 的维度使得下列和及乘积都有定义.

(a) $\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C}$　　（乘法的结合律）

(b) $\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C}) = \boldsymbol{AB}+\boldsymbol{AC}$　　（左分配律）

(c) $(\boldsymbol{B}+\boldsymbol{C})\boldsymbol{A} = \boldsymbol{BA}+\boldsymbol{CA}$　　（右分配律）

(d) 对任意数量 r，有 $r(\boldsymbol{AB}) = (r\boldsymbol{A})\boldsymbol{B} = \boldsymbol{A}(r\boldsymbol{B})$

(e) $\boldsymbol{I}_m\boldsymbol{A} = \boldsymbol{A} = \boldsymbol{AI}_n$　　（矩阵乘法的单位元）

证明： 性质(b)～(e)留作习题. 性质(a)可由矩阵乘法对应于线性变换(函数)的复合而得，我们知道函数的复合满足结合律(易证). 下面是性质(a)的另一种证明，它基于两矩阵乘积的“列定义”. 令

$$\boldsymbol{C}=[\boldsymbol{c}_1 \quad \cdots \quad \boldsymbol{c}_p]$$

由矩阵乘法的定义，

$$\boldsymbol{BC}=[\boldsymbol{Bc}_1 \quad \cdots \quad \boldsymbol{Bc}_p]$$
$$\boldsymbol{A}(\boldsymbol{BC})=[\boldsymbol{A}(\boldsymbol{Bc}_1) \quad \cdots \quad \boldsymbol{A}(\boldsymbol{Bc}_p)]$$

回顾(1)，$\boldsymbol{AB}$ 的定义使得对一切 $\boldsymbol{x}$ 都有 $\boldsymbol{A}(\boldsymbol{Bx})=(\boldsymbol{AB})\boldsymbol{x}$，所以

$$\boldsymbol{A}(\boldsymbol{BC})=[(\boldsymbol{AB})\boldsymbol{c}_1 \quad \cdots \quad (\boldsymbol{AB})\boldsymbol{c}_p]=(\boldsymbol{AB})\boldsymbol{C}$$ ■

定理1和2中的结合律、分配律表明，矩阵表达式中的括号可以像在实代数中那样插入或删除. 特别地，我们可以把 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 的乘积写成 $\boldsymbol{ABC}$，它既可以按 $\boldsymbol{A}(\boldsymbol{BC})$ 计算，也可以按 $(\boldsymbol{AB})\boldsymbol{C}$ 计算.[1] 类似地，四个矩阵的乘积 $\boldsymbol{ABCD}$ 可以按 $\boldsymbol{A}(\boldsymbol{BCD})$，$(\boldsymbol{ABC})\boldsymbol{D}$，$\boldsymbol{A}(\boldsymbol{BC})\boldsymbol{D}$ 等方式计算. 在计算矩阵乘积时，我们可以对矩阵任意分组，只要保持矩阵从左到右的顺序即可.

矩阵乘积中从左到右的顺序很重要，因为 $\boldsymbol{AB}$ 和 $\boldsymbol{BA}$ 通常不相等. 这也不足为奇，因为 $\boldsymbol{AB}$ 的列是 $\boldsymbol{A}$ 中列的线性组合，而 $\boldsymbol{BA}$ 的列是 $\boldsymbol{B}$ 中列的线性组合. 为了强调矩阵乘积中因子的顺序，称 $\boldsymbol{AB}$ 为 $\boldsymbol{A}$ *右乘以* $\boldsymbol{B}$ 或 $\boldsymbol{B}$ *左乘以* $\boldsymbol{A}$. 如果 $\boldsymbol{AB}=\boldsymbol{BA}$，我们说 $\boldsymbol{A}$，$\boldsymbol{B}$ 互相**可交换**(commute).

113

【例题7】 设 $\boldsymbol{A}=\begin{bmatrix}5 & 1\\ 3 & -2\end{bmatrix}$，$\boldsymbol{B}=\begin{bmatrix}2 & 0\\ 4 & 3\end{bmatrix}$，证明这两个矩阵不可交换，即，验证 $\boldsymbol{AB}\neq\boldsymbol{BA}$.

解：

$$\boldsymbol{AB}=\begin{bmatrix}5 & 1\\ 3 & -2\end{bmatrix}\begin{bmatrix}2 & 0\\ 4 & 3\end{bmatrix}=\begin{bmatrix}14 & 3\\ -2 & -6\end{bmatrix}$$

$$\boldsymbol{BA}=\begin{bmatrix}2 & 0\\ 4 & 3\end{bmatrix}\begin{bmatrix}5 & 1\\ 3 & -2\end{bmatrix}=\begin{bmatrix}10 & 2\\ 29 & -2\end{bmatrix}$$ ■

以下，我们列举了矩阵代数和普通实数代数中交换性的几点重要区别，以示强调. 具体例子见习题9～12.

警告：

1. 一般地，$\boldsymbol{AB}\neq\boldsymbol{BA}$.
2. 矩阵乘法中消去律不成立. 即，$\boldsymbol{AB}=\boldsymbol{AC}$ 时，$\boldsymbol{B}=\boldsymbol{C}$ 一般不成立. (见习题10.)
3. 如果 $\boldsymbol{AB}$ 是零矩阵，一般不能断定 $\boldsymbol{A}=\boldsymbol{0}$ 或 $\boldsymbol{B}=\boldsymbol{0}$. (见习题12.)

2.1.4 矩阵的幂

如果 $\boldsymbol{A}$ 是 $n\times n$ 矩阵，k 是正整数，则以 $\boldsymbol{A}^k$ 记 $\boldsymbol{A}$ 的 k 次幂：

$$\boldsymbol{A}^k=\underbrace{\boldsymbol{A}\cdots\boldsymbol{A}}_{k}$$

---

1. 当 $\boldsymbol{B}$ 是方阵且 $\boldsymbol{C}$ 的列数小于 $\boldsymbol{A}$ 的行数时，计算 $\boldsymbol{A}(\boldsymbol{BC})$ 比计算 $(\boldsymbol{AB})\boldsymbol{C}$ 效率更高一些.

如果 A 非零且 $x \in \mathbf{R}^n$，则 $A^k x$ 是 x 左乘 k 次 A 得到的．如果 $k=0$，$A^0 x$ 就是 x 本身．所以 A^0 解释成单位矩阵．矩阵的幂在理论和应用上都很有用（见2.6节和4.9节及本书后面的章节）．

2.1.5　矩阵的转置

给定 A 是 $m \times n$ 矩阵，A 的**转置**（transpose）是 $n \times m$ 矩阵，记为 A^{T}，它的列是 A 中对应的行．

【例题8】　设

114

$$A=\begin{bmatrix} a & b \\ c & d \end{bmatrix},\quad B=\begin{bmatrix} -5 & 2 \\ 1 & -3 \\ 0 & 4 \end{bmatrix},\quad C=\begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 5 & -2 & 7 \end{bmatrix}$$

则

$$A^{\mathrm{T}}=\begin{bmatrix} a & c \\ b & d \end{bmatrix},\quad B^{\mathrm{T}}=\begin{bmatrix} -5 & 1 & 0 \\ 2 & -3 & 4 \end{bmatrix},\quad C^{\mathrm{T}}=\begin{bmatrix} 1 & -3 \\ 1 & 5 \\ 1 & -2 \\ 1 & 7 \end{bmatrix}$$ ■

【定理3】　设 A，B 是使下列和及乘积有定义的矩阵，

(a) $(A^{\mathrm{T}})^{\mathrm{T}}=A$

(b) $(A+B)^{\mathrm{T}}=A^{\mathrm{T}}+B^{\mathrm{T}}$

(c) 对任数量 r，$(rA)^{\mathrm{T}}=rA^{\mathrm{T}}$

(d) $(AB)^{\mathrm{T}}=B^{\mathrm{T}}A^{\mathrm{T}}$

(a)～(c)的证明很简单，在此省略．(d)的证明见习题33．即便 A 和 B 的维度使 $A^{\mathrm{T}}B^{\mathrm{T}}$ 有定义，$(AB)^{\mathrm{T}}$ 与 $A^{\mathrm{T}}B^{\mathrm{T}}$ 一般来说并不相等．

定理3中的(d)推广到两个以上矩阵的乘积时，可以表述为：

矩阵乘积的转置等于它们转置的逆序乘积．

矩阵转置的上述性质在本节习题，特别是计算题中都有体现．

注记

1. 在计算机上用哪种方法计算 AB 速度最快？这取决于矩阵在计算机中的存储方式．标准的高性能算法，比如LAPACK，与我们这里所定义的乘积一样按列计算 AB．（LAPACK的一个C++编写的版本则按行计算 AB．）

2. AB 的定义使得在计算机上并行计算 AB 十分容易．AB 的列被单个或成组地分配给不同的处理器，这些处理器彼此独立工作，同步计算 AB 中对应的列．

115

基础练习

1. 因为 $\mathbf{R}^n$ 中的向量可以看成 $n \times 1$ 矩阵，所以定理3中的转置性质也可以应用到向量上．设

$$A=\begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix} \text{且} x=\begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

计算 $(Ax)^{\mathrm{T}}$，$x^{\mathrm{T}}A^{\mathrm{T}}$，$xx^{\mathrm{T}}$，$x^{\mathrm{T}}x$．$A^{\mathrm{T}}x^{\mathrm{T}}$ 有定义吗？

2. 设 A 是 4×4 矩阵，x 是 $\mathbf{R}^4$ 中的向量．计算 A^2x 的最快方法是什么？乘法运算的次数是

多少?

习题 2.1

在习题 1 和 2 中，如果矩阵的和或者乘积有意义，计算其结果，否则说明原因. 其中，

$$A=\begin{bmatrix}2&0&-1\\4&-5&2\end{bmatrix},\ B=\begin{bmatrix}7&-5&1\\1&-4&-3\end{bmatrix},\ C=\begin{bmatrix}1&2\\-2&1\end{bmatrix},\ D=\begin{bmatrix}3&5\\-1&4\end{bmatrix},\ E=\begin{bmatrix}-5\\3\end{bmatrix}$$

1. $-2A$, $B-2A$, AC, CD
2. $A+2B$, $3C-E$, CB, EB

在下面所有习题中，假定每个矩阵表达式都有意义，即矩阵(还有向量)的维度都恰好匹配.

3. 设 $A=\begin{bmatrix}4&-1\\5&-2\end{bmatrix}$，计算 $3I_2-A$，$(3I_2)A$.
4. 计算 $A-5I_3$ 和 $(5I_3)A$，其中

$$A=\begin{bmatrix}9&-1&3\\-8&7&-6\\-4&1&8\end{bmatrix}.$$

在习题 5 和 6 中，用两种方法计算乘积 AB：(a)用定义，其中 Ab_1 和 Ab_2 分开计算，(b)用计算 AB 的行列法则.

5. $A=\begin{bmatrix}-1&2\\5&4\\2&-3\end{bmatrix}$，$B=\begin{bmatrix}3&-2\\-2&1\end{bmatrix}$　　6. $A=\begin{bmatrix}4&-2\\-3&0\\3&5\end{bmatrix}$，$B=\begin{bmatrix}1&3\\2&-1\end{bmatrix}$
7. 如果 A 是 5×3 矩阵，乘积 AB 是 5×7 矩阵，则 B 的维度是多少?
8. 如果 BC 是 3×4 矩阵，则 B 的行数是多少?
9. 设 $A=\begin{bmatrix}2&5\\-3&1\end{bmatrix}$，$B=\begin{bmatrix}4&-5\\3&k\end{bmatrix}$，则 k 取何值时 $AB=BA$.
10. 设 $A=\begin{bmatrix}2&-3\\-4&6\end{bmatrix}$，$B=\begin{bmatrix}8&4\\5&5\end{bmatrix}$，$C=\begin{bmatrix}5&-2\\3&1\end{bmatrix}$. 证明 $AB=AC$ 但 $B\neq C$.
11. 设 $A=\begin{bmatrix}1&1&1\\1&2&3\\1&4&5\end{bmatrix}$，$D=\begin{bmatrix}2&0&0\\0&3&0\\0&0&5\end{bmatrix}$. 计算 AD 和 DA. 解释用 D 右乘或左乘 A 时 A 的列和行将怎样变化. 求一个非单位、非零的 3×3 矩阵 B，使得 $AB=BA$.
12. 设 $A=\begin{bmatrix}3&-6\\-1&2\end{bmatrix}$，构造一个 2×2 矩阵 B 使得 AB 是零矩阵. 要求 B 的两列非零且不相等.
13. 设 $r_1,\cdots,r_p$ 是 $\mathbb{R}^n$ 中的向量，Q 是一个 $m\times n$ 矩阵. 将矩阵 $[Qr_1\cdots Qr_p]$ 写成两个矩阵(两个都非单位矩阵)的乘积.
14. 设 U 是 1.8 节例题 6 所述的 3×2 成本矩阵. U 的第一、二列分别列出了生产产品 **B** 和 **C** 的“一美元产出成本”(成本分为原材料、劳动力和管理费). 设 q_1，q_2，q_3 和 q_4 是 $\mathbb{R}^2$ 中的向量，它们依次列出了本年度四个季度内产品 **B** 和 **C** 的产量(以美元度量). 试解释矩阵

UQ 中元素的经济含义，其中 $Q = [q_1 \quad q_2 \quad q_3 \quad q_4]$.

习题 15 和 16 中的任意矩阵 A，B，C 使得下面的和与乘积都有意义. 判断下列各命题的真假，并说明理由.

15. a. 如果 A，B 都是 2×2 矩阵，其列向量分别为 a_1，a_2 和 b_1，b_2，则 $AB = [a_1b_1 \quad a_2b_2]$.
    b. AB 的每一列都是 B 的列的线性组合，权重为 A 的相应列.
    c. $AB + AC = A(B + C)$.
    d. $A^T + B^T = (A + B)^T$.
    e. 矩阵乘积的转置等于它们转置在同一顺序下的乘积.
16. a. 如果 A，B 是 3×3 矩阵，$B = [b_1 \quad b_2 \quad b_3]$，则 $AB = [Ab_1 + Ab_2 + Ab_3]$.
    b. AB 的第二行是由 A 的第二行右乘 B 得到.
    c. $(AB)C = (AC)B$
    d. $(AB)^T = A^TB^T$
    e. 矩阵和的转置等于它们转置的和.
17. 如果 $A = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}$，$AB = \begin{bmatrix} -1 & 2 & -1 \\ 6 & -9 & 3 \end{bmatrix}$，确定 B 的第一、二列.
18. 假设 B 的前两列 b_1 和 b_2 相等，则对 AB 的列将有什么结论(如果 AB 有意义)？为什么？
19. 假设 B 的第三列是前两列的和，则对 AB 的第三列将有什么结论？为什么？
20. 假设 B 的第二列全为零，则对 AB 的第二列将有什么结论？
21. 假设 AB 的最后一列全为零，但 B 本身没有零列，则对 A 的列将有什么结论？
22. 证明：如果 B 的列向量线性相关，则 AB 的列向量也线性相关.
23. 假设 $CA = I_n$($n\times n$ 单位矩阵)，证明：方程 $Ax = 0$ 只有零解. 解释 A 的列数为什么不能比行数多.
24. 假设 $AD = I_m$($m\times m$ 单位矩阵)，证明：对 $\mathbf{R}^m$ 中任意向量 b，方程 $Ax = b$ 只有唯一解. [提示：考虑方程 $ADb = b$.]解释 A 的行数为什么不能比列数多.
25. 假设 A 是 $m\times n$ 矩阵，存在 $n\times m$ 矩阵 C，D 使得 $CA = I_n$ 且 $AD = I_m$. 证明：$m = n$ 且 $C = D$. [提示：考虑乘积 CAD.]
26. 假设 A 是一个 $3\times n$ 矩阵，它的列张成 $\mathbf{R}^3$. 解释怎样构造一个 $n\times 3$ 矩阵使得 $AD = I_3$.

在习题 27 和 28 中，将 $\mathbf{R}^n$ 中的向量看成 $n\times 1$ 矩阵. 对 $\mathbf{R}^n$ 中的向量 u、v，矩阵乘积 u^Tv 是 1×1 矩阵，称为 u 和 v 的**数量积**(scalar product)或**内积**(inner product)，通常写成一个单独的无括号的实数. 矩阵乘积 uv^T 是 $n\times n$ 矩阵，称为 u 和 v 的**外积**(outer product). 矩阵乘积 u^Tv 和 uv^T 将在本书的后面出现.

27. 设 $u = \begin{bmatrix} -2 \\ 3 \\ -4 \end{bmatrix}$，$v = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. 计算 u^Tv，v^Tu，uv^T 和 vu^T.
28. 如果 u，v 是 $\mathbf{R}^n$ 中的向量，则 u^Tv 和 v^Tu 有什么关系？uv^T 和 vu^T 有什么关系？
29. 用行列法则证明定理 2(b) 和 2(c)，$A(B + C)$ 中的 (i,j) - 元可以写成
$$a_{i1}(b_{1j} + c_{1j}) + \cdots + a_{in}(b_{nj} + c_{nj}) \text{或} \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj})$$
30. 证明定理 2(d). [提示：$(rA)B$ 中的 (i,j) - 元是 $(ra_{i1})b_{1j} + \cdots + (ra_{in})b_{nj}$.]
31. A 是 $m\times n$ 矩阵，证明 $I_mA = A$. 可以假设对 $\mathbf{R}^m$ 中任意向量 x 有 $I_mx = x$.
32. A 是 $m\times n$ 矩阵，证明 $AI_n = A$. [提示：用 AI_n 中列的定义.]
33. 证明定理 3(d). [提示：考虑 $(AB)^T$ 的第 j 行.]

34. 给出$(ABx)^T$的一个公式，其中x是向量，A，B是有适当维度的矩阵.
35. [M]阅读矩阵软件的相关文档，写出能生成(自动生成，不需要键入矩阵的每一个元素)下列矩阵的命令：
   a. 元素全为零的5×6矩阵.　b. 元素全为1的3×5矩阵.
   c. 6×6单位矩阵.　d. 5×5对角矩阵，对角线元素是3、5、7、2、4.

   在矩阵代数中，为了验证新的命题或者进行某种猜想，一个常用的方法是用随机矩阵进行演算. 证明某个性质对某些矩阵成立，尽管不能说明该性质普遍成立，但是可以让我们更有把握. 另外，如果这个性质确实是错的，那么通过举例演算也许可以发现.
36. [M]写出能够生成6×4随机矩阵(其中的元素都是随机数)的命令. 这些矩阵元素的取值范围是什么？说出如何生成一个3×3矩阵，使得其中的随机整数元在-9到9之间. [提示：如果x是一个满足$0<x<1$的随机数，则$-9.5<19(x-0.5)<9.5$.]

117

37. [M]构造一个4×4随机矩阵A并检验是否有$(A+I)(A-I)=A^2-I$. 最好的方式是计算$(A+I)(A-I)-(A^2-I)$，并验证这个差是零矩阵. 任选三个随机矩阵验证这一结论. 接下来，任选三对4×4随机矩阵，利用同样的方法验证$(A+B)(A-B)=A^2-B^2$. 报告你的结论.
38. [M]至少利用三对4×4随机矩阵A和B来检验等式$(A+B)^T=A^T+B^T$和$(AB)^T=B^TA^T$(见习题37.). 给出你的结论. [注意：大部分矩阵软件中以A'代替A^T.]
39. [M]设

$$S=\begin{bmatrix}0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\\0&0&0&0&0\end{bmatrix}$$

   对于$k=2,\cdots,6$，计算S^k.
40. [M]设

$$A=\begin{bmatrix}1/6&1/2&1/3\\1/2&1/4&1/4\\1/3&1/4&5/12\end{bmatrix}$$

   用文字描述计算A^5，A^{10}，A^{20}和A^{30}时会出现什么问题？

基础练习答案

1. $Ax=\begin{bmatrix}1&-3\\-2&4\end{bmatrix}\begin{bmatrix}5\\3\end{bmatrix}=\begin{bmatrix}-4\\2\end{bmatrix}$. 因此$(Ax)^T=[-4\quad 2]$，$x^TA^T=[5\quad 3]\begin{bmatrix}1&-2\\-3&4\end{bmatrix}=[-4\quad 2]$. 如定理3(d)所述，$(Ax)^T$与$x^TA^T$相等. 此外，

$$xx^T=\begin{bmatrix}5\\3\end{bmatrix}[5\quad 3]=\begin{bmatrix}25&15\\15&9\end{bmatrix}$$

$$x^Tx=[5\quad 3]\begin{bmatrix}5\\3\end{bmatrix}=[25+9]=34$$

   1×1矩阵，比如x^Tx，在书写时通常不加方括号. 由于x^T的行数与A^T的列数不等，所以A^Tx^T无意义.
2. 计算A^2x最快的方法是计算$A(Ax)$. 计算乘积Ax需要做16次乘法，其中每个元素4次，$A(Ax)$则在Ax的基础上再做16次. 而计算乘积A^2需要做64次乘法，A^2中16个元素每个需要4次. 求出A^2之后，计算A^2x还需16次乘法，总共80次.

2.2 矩阵的逆

类似于普通的实代数，矩阵代数可以帮助我们处理矩阵方程并且建立许多常用公式. 本节将非零实数的倒数或称乘法逆元类推到矩阵上，讨论矩阵的逆.

回顾一下实数的乘法逆元，例如5的乘法逆元是1/5或5^{-1}. 这个逆元满足方程

118

$$5^{-1}\cdot 5=1 \text{ 且 } 5\cdot 5^{-1}=1$$

由于矩阵乘法不满足交换律，因此将逆元概念推广到矩阵时，上述两个方程需同时满足，并且此时应避免使用斜线记号(表示除). 此外，仅当我们所讨论的矩阵都是方阵时，才有可能得到一个完全的推广.[1]

2.2.1 矩阵逆的概念

称$n\times n$矩阵A是**可逆**(invertible)的，如果存在一个$n\times n$矩阵C使得

$$CA=I \text{ 且 } AC=I$$

其中$I=I_n$是$n\times n$单位矩阵. 这时，C称作A的**逆**(inverse). 实际上，C由A唯一确定，因为如果B是A的另一个逆，则$B=BI=B(AC)=(BA)C=IC=C$. 将这个唯一的逆记为A^{-1}，于是

$$A^{-1}A=I \text{ 且 } AA^{-1}=I$$

不可逆矩阵和可逆矩阵也分别称为**奇异矩阵**(singular matrix)和**非奇异矩阵**(nonsingular matrix).

【例题1】 设$A=\begin{bmatrix}2&5\\-3&-7\end{bmatrix}$，$C=\begin{bmatrix}-7&-5\\3&2\end{bmatrix}$，则

$$AC=\begin{bmatrix}2&5\\-3&-7\end{bmatrix}\begin{bmatrix}-7&-5\\3&2\end{bmatrix}=\begin{bmatrix}1&0\\0&1\end{bmatrix}$$

$$CA=\begin{bmatrix}-7&-5\\3&2\end{bmatrix}\begin{bmatrix}2&5\\-3&-7\end{bmatrix}=\begin{bmatrix}1&0\\0&1\end{bmatrix}$$

所以$C=A^{-1}$. ■

下面给出了2×2矩阵的求逆公式，它同时指明了逆矩阵存在的条件.

【定理4】 设$A=\begin{bmatrix}a&b\\c&d\end{bmatrix}$. 如果$ad-bc\neq 0$，则$A$是可逆的且

$$A^{-1}=\frac{1}{ad-bc}\begin{bmatrix}d&-b\\-c&a\end{bmatrix}$$

如果$ad-bc=0$，则A是不可逆的.

在习题25和26中概述了定理4的证明. 数量$ad-bc$称为A的**行列式**(determinant)，记为

---

1. 称$m\times n$矩阵A是可逆的，如果存在$n\times m$矩阵C和D使得$CA=I_n$，$AD=I_m$. 不过，这些方程表明A是方阵并且$C=D$. 因此A是可逆的. 见2.1节习题23~25.

$$\det A = ad - bc$$

定理4表明，2×2矩阵A是可逆的当且仅当 $\det A \neq 0$.

119

【例题2】 求 $A=\begin{bmatrix}3 & 4\\5 & 6\end{bmatrix}$ 的逆矩阵.

解： 因为 $\det A = 3(6)-4(5) = -2 \neq 0$，所以 A 是可逆的，且

$$A^{-1}=\frac{1}{-2}\begin{bmatrix}6 & -4\\-5 & 3\end{bmatrix}=\begin{bmatrix}6/(-2) & -4/(-2)\\-5/(-2) & 3/(-2)\end{bmatrix}=\begin{bmatrix}-3 & 2\\5/2 & -3/2\end{bmatrix}$$ ■

可逆矩阵的概念在线性代数中不可或缺，它主要应用于代数计算和公式推导，比如下面的定理. 还有一些场合，比如在下面的例题3中，逆矩阵有助于我们理解实际问题中的数学模型.

【定理5】 如果A是 $n\times n$ 可逆矩阵，则对任意 $\boldsymbol{b}\in\mathbb{R}^n$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 都有唯一解 $\boldsymbol{x}=A^{-1}\boldsymbol{b}$.

证明： 任取 $b\in\mathbb{R}^n$. 因为如果用 $A^{-1}\boldsymbol{b}$ 代替 $\boldsymbol{x}$，则 $A\boldsymbol{x}=A(A^{-1}\boldsymbol{b})=(AA^{-1})\boldsymbol{b}=I\boldsymbol{b}=\boldsymbol{b}$，所以解是存在的，$A^{-1}\boldsymbol{b}$ 就是一个解. 为了证明解的唯一性，设 $\boldsymbol{u}$ 是任意一个解，则需证明 $\boldsymbol{u}$ 就是 $A^{-1}\boldsymbol{b}$. 实际上，如果有 $A\boldsymbol{u}=\boldsymbol{b}$，我们在两端同乘以 A^{-1} 得到

$$A^{-1}A\boldsymbol{u}=A^{-1}\boldsymbol{b},\ I\boldsymbol{u}=A^{-1}\boldsymbol{b},\ \boldsymbol{u}=A^{-1}\boldsymbol{b}$$ ■

【例题3】 如图2-5所示，水平弹性横梁的支撑点位于两端，且在点1、2、3处被施加了作用力. 设 $\boldsymbol{f}\in\mathbb{R}^3$ 列出了施加在这三个点上的作用力，$\boldsymbol{y}\in\mathbb{R}^3$ 列出了横梁在这三个点的偏移(移动). 利用物理学中的胡克定律，可以证明

$$\boldsymbol{y}=D\boldsymbol{f}$$

其中 D 是柔度矩阵，其逆矩阵称为刚度矩阵. 试描述矩阵 D 和 D^{-1} 中列向量的物理意义.

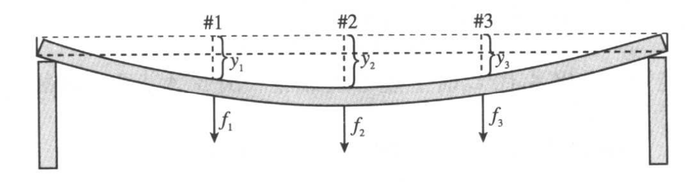


图2-5 弹性横梁的偏移

解： 记 $I_3=[\boldsymbol{e}_1\quad \boldsymbol{e}_2\quad \boldsymbol{e}_3]$，观察

$$D=DI_3=[D\boldsymbol{e}_1\quad D\boldsymbol{e}_2\quad D\boldsymbol{e}_3]$$

120

将向量 $\boldsymbol{e}_1=(1,0,0)$ 解释为在横梁点1处施加向下的1单位作用力(而其余两点处的作用力为0). 则 $D\boldsymbol{e}_1$，即 D 的第一列，列出了由点1处的1单位作用力所引起的横梁偏移. D 的第二、三列也有类似的解释.

为了研究刚度矩阵 D^{-1}，注意到，当偏移向量 $\boldsymbol{y}$ 给定时，由方程 $\boldsymbol{f}=D^{-1}\boldsymbol{y}$ 可计算出作用力向量 $\boldsymbol{f}$. 记

$$D^{-1}=D^{-1}I_3=[D^{-1}\boldsymbol{e}_1\quad D^{-1}\boldsymbol{e}_2\quad D^{-1}\boldsymbol{e}_3]$$

将e_1解释成偏移向量，则$D^{-1}e_1$列出了产生这个偏移的作用力．即D^{-1}的第一列列出了横梁三个点处的作用力，它使点1产生1单位偏移，其余两点的偏移为零．类似地，D^{-1}的第二、三列分别列出了在点2和点3产生1单位偏移的作用力．每一列中，为使横梁在某点产生1单位偏移在另两点不偏移，必须有一个或两个力是负(向上)的．如果柔度以英寸/磅为单位(其中英寸和磅分别是偏移和作用力的单位)，则刚度矩阵中元素的单位是磅/英寸．■

定理5中的公式很少用于方程$Ax=b$的数值求解，因为$[A\quad b]$的行化简会更快．(当计算中包含舍入误差时，行化简通常也更准确．)2×2矩阵可能是一个例外，这时，利用A^{-1}的公式心算求解$Ax=b$有时更容易，比如下面这道例题．

【例题4】　利用例题2中矩阵A的逆解方程组

$$3x_1+4x_2=3$$
$$5x_1+6x_2=7$$

解：此方程组等价于$Ax=b$，所以

$$x=A^{-1}b=\begin{bmatrix}-3 & 2\\5/2 & -3/2\end{bmatrix}\begin{bmatrix}3\\7\end{bmatrix}=\begin{bmatrix}5\\-3\end{bmatrix}.$$ ■

下面的定理给出了三个关于逆矩阵的有用结论．

【定理6】

a. 如果A是可逆矩阵，则A^{-1}也是可逆的，且

$$(A^{-1})^{-1}=A$$

b. 如果A和B是$n\times n$可逆矩阵，则AB也是$n\times n$可逆矩阵，且AB的逆是A的逆与B的逆的逆序乘积．即

$$(AB)^{-1}=B^{-1}A^{-1}$$

c. 如果A是可逆矩阵，则A^T也是可逆的，且A^T的逆是A^{-1}的转置．即

$$(A^T)^{-1}=(A^{-1})^T$$

121

证明：为了证明(a)，我们必须找一个矩阵C使得

$$A^{-1}C=I \text{ 且 } CA^{-1}=I$$

然而，我们已经知道，以A代C时上述等式成立．因此A^{-1}是可逆的，A就是它的逆．为了证明(b)，计算：

$$(AB)(B^{-1}A^{-1})=A(BB^{-1})A^{-1}=AIA^{-1}=AA^{-1}=I$$

类似可得$(B^{-1}A^{-1})(AB)=I$．对于(c)，利用定理3(d)(从右向左看)，$(A^{-1})^TA^T=(AA^{-1})^T=I^T=I$．类似地，$A^T(A^{-1})^T=I^T=I$．因此$A^T$是可逆的，$(A^{-1})^T$就是它的逆．■

下面的命题是定理6(b)的推广，后面将要用到．

$n\times n$可逆矩阵的乘积是可逆的，它的逆是这些矩阵的逆的逆序的乘积．

可逆矩阵与行变换之间有一个重要的联系，它将推导出求逆矩阵的一种方法．正如我们下面将看到的，可逆矩阵A等价于单位矩阵，通过观察从A到I的行化简我们可以求出A^{-1}．

2.2.2　初等矩阵

初等矩阵(elementary matrix)是在单位矩阵上施行一次初等行变换得到的．下面

的例子列举了三种初等矩阵.

【例题 5】 设

$$E_1=\begin{bmatrix}1&0&0\\0&1&0\\-4&0&1\end{bmatrix},\ E_2=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix},\ E_3=\begin{bmatrix}1&0&0\\0&1&0\\0&0&5\end{bmatrix},\ A=\begin{bmatrix}a&b&c\\d&e&f\\g&h&i\end{bmatrix}$$

计算 E_1A，E_2A 和 E_3A，试描述这些乘积怎样由 A 上的初等行变换得到.

解： 我们有

$$E_1A=\begin{bmatrix}a&b&c\\d&e&f\\g-4a&h-4b&i-4c\end{bmatrix},\ E_2A=\begin{bmatrix}d&e&f\\a&b&c\\g&h&i\end{bmatrix},\ E_3A=\begin{bmatrix}a&b&c\\d&e&f\\5g&5h&5i\end{bmatrix}$$

122

A 的第一行乘以 -4 加到第三行得到 E_1A（这是一个行替换运算）. 交换 A 的第一、二行得到 E_2A，A 的第三行乘以 5 得到 E_3A. ■

对任意 $3\times n$ 矩阵左乘（即在左边乘）例题 5 中的 E_1 也会产生同样的效果，即把第一行的 -4 倍加到第三行上. 特别地，因为 $E_1\cdot I=E_1$，可见 E_1 本身也是由单位矩阵经同一行变换而得到的. 所以，例题 5 说明了关于初等矩阵的下述一般事实，见习题 27 和 28.

如果在 $m\times n$ 矩阵 A 上施加一个初等行变换，得到的矩阵可以写成 EA，其中 $m\times m$ 矩阵 E 是在 I_m 上施加同一行变换而得到的矩阵.

如 1.1 节所示，行变换是可逆的，所以初等矩阵也是可逆的：如果 E 由 I 施加一个行变换得到，则存在同一类型的另一个行变换把 E 变回到 I. 因此存在一个初等矩阵 F 使得 $FE=I$. 由于 E 和 F 分别对应着互逆的两个变换，所以也有 $EF=I$.

每一个初等矩阵 E 都是可逆的，E 的逆是与 E 属同一类型的初等矩阵，它将 E 变回到 I.

【例题 6】 求 $E_1=\begin{bmatrix}1&0&0\\0&1&0\\-4&0&1\end{bmatrix}$的逆矩阵.

解： 为了将 E_1 变换到 I，将第一行的 $+4$ 倍加到第三行. 此操作对应的初等矩阵是

$$E_1^{-1}=\begin{bmatrix}1&0&0\\0&1&0\\+4&0&1\end{bmatrix}$$ ■

下列定理给出了矩阵可逆的最佳判别法，利用它还可以立即求出矩阵的逆.

【定理 7】 **$n\times n$ 矩阵 A 是可逆的当且仅当 A 行等价于 I_n. 当 A 可逆时，将 A 化简到 I_n 的任意一个初等行变换序列同时也将 I_n 变换到 A^{-1}.**

123

证明： 假设 A 是可逆的. 由于对每个 b 方程 $Ax=b$ 都有解（定理 5），所以 A 的每一行上都有一个主元位置（1.4 节定理 4）. 因为 A 是方阵，这 n 个主元位置一定在对角线上，这就表明 A 的简化阶梯形式是 I_n，即 $A\sim I_n$.

反过来，假设 $A\sim I_n$. 因为 A 的行化简的每一步都对应着左乘以一个初等矩阵，所以存在初等矩阵 $E_1,\cdots,E_p$ 使得

$$A \sim E_1A \sim E_2(E_1A) \sim \cdots \sim E_p(E_{p-1}\cdots E_1A) = I_n$$

即，

$$E_p\cdots E_1A = I_n \tag{1}$$

因为可逆矩阵的乘积 $E_p\cdots E_1$ 也是可逆的，由(1)有

$$(E_p\cdots E_1)^{-1}(E_p\cdots E_1)A = (E_p\cdots E_1)^{-1}I_n$$

$$A = (E_p\cdots E_1)^{-1}$$

A 是一个可逆矩阵的逆，因此 A 是可逆的(定理6)．此外还有，

$$A^{-1} = [(E_p\cdots E_1)^{-1}]^{-1} = E_p\cdots E_1$$

于是 $A^{-1} = E_p\cdots E_1\cdot I_n$，这说明 A^{-1} 是由 $E_1,\cdots,E_p$ 连续作用到 I_n 上得到的．该序列与(1)中将 A 化简为 I_n 的序列相同． ■

2.2.3　求 A^{-1} 的算法

如果我们把 A 和 I 并列排成一个增广矩阵 $[A\quad I]$，那么该矩阵上的行变换将在 A 和 I 上施加相同的变换．根据定理7，或者存在将 A 变换到 I_n 且将 I_n 变换到 A^{-1} 的行变换，或者 A 不可逆．

求 A^{-1} 的算法

行化简增广矩阵 $[A\quad I]$．如果 A 行等价于 I，则 $[A\quad I]$ 行等价于 $[I\quad A^{-1}]$．否则，A 没有逆矩阵．

【例题7】　若矩阵 $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$ 可逆，求其逆矩阵．

解：

$$[A\quad I] = \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix}$$

因为 $A \sim I$，由定理7我们断定 A 可逆，且

$$A^{-1} = \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix}$$

最好检验一下最后结果：

$$AA^{-1} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}\begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

124

由于 $\boldsymbol{A}$ 可逆，因此无需再验证 $\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{I}$. ■

2.2.4 矩阵求逆的另一种见解

记 $\boldsymbol{I}_n$ 的列向量为 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n$. 则 $[\boldsymbol{A}\quad \boldsymbol{I}]$ 到 $[\boldsymbol{I}\quad \boldsymbol{A}^{-1}]$ 的行化简可以看成下面 n 个方程组的公共解：

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{e}_1,\boldsymbol{A}\boldsymbol{x}=\boldsymbol{e}_2,\cdots,\boldsymbol{A}\boldsymbol{x}=\boldsymbol{e}_n \tag{2}$$

将这些方程组的“增广列”都列于 $\boldsymbol{A}$ 之后就得到 $[\boldsymbol{A}\quad \boldsymbol{e}_1\quad \boldsymbol{e}_2\quad \cdots \boldsymbol{e}_n]=[\boldsymbol{A}\quad \boldsymbol{I}]$. 方程 $\boldsymbol{A}\boldsymbol{A}^{-1}=\boldsymbol{I}$ 以及矩阵乘法的定义表明，$\boldsymbol{A}^{-1}$ 的列向量恰好是(2)中方程组的解. 这个发现很有用，因为一些实际问题中可能只需要求出 $\boldsymbol{A}^{-1}$ 的一两列，此时求解(2)中对应的方程组即可.

注记

在实际工作中，很少计算 $\boldsymbol{A}^{-1}$，除非我们需要得到 $\boldsymbol{A}^{-1}$ 中的全部元素. 同时计算出 $\boldsymbol{A}^{-1}$ 和 $\boldsymbol{A}^{-1}\boldsymbol{b}$ 所需的运算量是用行化简方法求解 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的三倍，且行化简可能更精确.

基础练习

1. 利用行列式确定下列哪些矩阵是可逆的.

a. $\begin{bmatrix}3 & -9\\ 2 & 6\end{bmatrix}$ b. $\begin{bmatrix}4 & -9\\ 0 & 5\end{bmatrix}$ c. $\begin{bmatrix}6 & -9\\ -4 & 6\end{bmatrix}$

2. 若矩阵 $\boldsymbol{A}=\begin{bmatrix}1 & -2 & -1\\ -1 & 5 & 6\\ 5 & -4 & 5\end{bmatrix}$ 可逆，求其逆矩阵.

125

习题 2.2

求习题 1 ~ 4 中矩阵的逆.

1. $\begin{bmatrix}8 & 6\\ 5 & 4\end{bmatrix}$ 2. $\begin{bmatrix}3 & 2\\ 7 & 4\end{bmatrix}$ 3. $\begin{bmatrix}8 & 5\\ -7 & -5\end{bmatrix}$ 4. $\begin{bmatrix}3 & -4\\ 7 & -8\end{bmatrix}$

5. 利用习题 1 中求得的逆矩阵解方程组

$8x_1+6x_2=2$

$5x_1+4x_2=-1$

6. 利用习题 3 中求得的逆矩阵解方程组

$8x_1+5x_2=-9$

$-7x_1-5x_2=11$

7. 设 $\boldsymbol{A}=\begin{bmatrix}1 & 2\\ 5 & 12\end{bmatrix}$，$\boldsymbol{b}_1=\begin{bmatrix}-1\\ 3\end{bmatrix}$，$\boldsymbol{b}_2=\begin{bmatrix}1\\ -5\end{bmatrix}$，$\boldsymbol{b}_3=\begin{bmatrix}2\\ 6\end{bmatrix}$ 且 $\boldsymbol{b}_4=\begin{bmatrix}3\\ 5\end{bmatrix}$.

a. 求 $\boldsymbol{A}^{-1}$，并且利用它求解下列四个方程：

$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}_1$， $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}_2$， $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}_3$， $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}_4$

b. 由于 a. 中四个方程有相同的系数矩阵，因此可以用同一个行变换集求解，即对增广矩阵 $[\boldsymbol{A}\quad \boldsymbol{b}_1\quad \boldsymbol{b}_2\quad \boldsymbol{b}_3\quad \boldsymbol{b}_4]$ 进行行化简可同时求出四个方程的解.

8. 利用矩阵代数说明如果 $\boldsymbol{A}$ 是可逆的，且 $\boldsymbol{D}$ 满足 $\boldsymbol{A}\boldsymbol{D}=\boldsymbol{I}$，则 $\boldsymbol{D}=\boldsymbol{A}^{-1}$.

在习题 9 和 10 中，判断下列各命题的真假，并说明理由.

9. a. 为了使 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的逆矩阵，方程 $\boldsymbol{AB}=\boldsymbol{I}$ 和 $\boldsymbol{BA}=\boldsymbol{I}$ 必须同时成立.
   b. 如果 $\boldsymbol{A}$，$\boldsymbol{B}$ 是 $n\times n$ 可逆矩阵，则 $\boldsymbol{A}^{-1}\boldsymbol{B}^{-1}$ 是 $\boldsymbol{AB}$ 的逆.
   c. 如果 $\boldsymbol{A}=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$，且 $ab-cd\neq 0$，则 $\boldsymbol{A}$ 是可逆的.
   d. 如果 $\boldsymbol{A}$ 是 $n\times n$ 可逆矩阵，则对任意 $\boldsymbol{b}\in\mathbf{R}^n$，方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 都是相容的.
   e. 每一个初等矩阵都是可逆的.
10. a. 若干个 $n\times n$ 可逆矩阵的乘积是可逆的，这个乘积的逆等于这些矩阵的逆在同一顺序下的乘积.
   b. 如果 $\boldsymbol{A}$ 可逆，则 $\boldsymbol{A}^{-1}$ 的逆是 $\boldsymbol{A}$ 本身.
   c. 如果 $\boldsymbol{A}=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$，且 $ad=bc$，则 $\boldsymbol{A}$ 是不可逆的.
   d. 如果 $\boldsymbol{A}$ 可以行化简为单位矩阵，则 $\boldsymbol{A}$ 必定是可逆的.
   e. 如果 $\boldsymbol{A}$ 是可逆的，则将 $\boldsymbol{A}$ 化简为 $\boldsymbol{I}_n$ 的初等行变换也将 $\boldsymbol{A}^{-1}$ 化简为 $\boldsymbol{I}_n$.
11. 设 $\boldsymbol{A}$ 是 $n\times n$ 可逆矩阵，$\boldsymbol{B}$ 是 $n\times p$ 可逆矩阵. 证明：方程 $\boldsymbol{AX}=\boldsymbol{B}$ 有唯一解 $\boldsymbol{A}^{-1}\boldsymbol{B}$.
12. 设 $\boldsymbol{A}$ 是 $n\times n$ 可逆矩阵，$\boldsymbol{B}$ 是 $n\times p$ 可逆矩阵. 解释为什么 $\boldsymbol{A}^{-1}\boldsymbol{B}$ 可由下面的行化简计算得到：

$$\text{如果}[\boldsymbol{A}\quad \boldsymbol{B}]\sim\cdots\sim[\boldsymbol{I}\quad \boldsymbol{X}]\text{，则 }\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}.$$

   如果 $\boldsymbol{A}$ 是大于 2×2 的矩阵，则对 $[\boldsymbol{A}\quad \boldsymbol{B}]$ 进行行化简比计算 $\boldsymbol{A}^{-1}$ 和 $\boldsymbol{A}^{-1}\boldsymbol{B}$ 快.
13. 假设 $\boldsymbol{AB}=\boldsymbol{AC}$，其中 $\boldsymbol{B}$，$\boldsymbol{C}$ 是 $n\times p$ 矩阵，$\boldsymbol{A}$ 是可逆的. 证明 $\boldsymbol{B}=\boldsymbol{C}$. 一般地，如果 $\boldsymbol{A}$ 是不可逆的，这个结论还对吗？
14. 假设 $(\boldsymbol{B}-\boldsymbol{C})\boldsymbol{D}=0$，其中 $\boldsymbol{B}$，$\boldsymbol{C}$ 是 $m\times n$ 矩阵，$\boldsymbol{D}$ 是可逆的. 证明 $\boldsymbol{B}=\boldsymbol{C}$.
15. 假设 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 是 $n\times n$ 可逆矩阵，通过找出一个矩阵 $\boldsymbol{D}$ 使得 $(\boldsymbol{ABC})\boldsymbol{D}=\boldsymbol{I}$ 和 $\boldsymbol{D}(\boldsymbol{ABC})=\boldsymbol{I}$ 来证明 $\boldsymbol{ABC}$ 是可逆的.
16. 假设矩阵 $\boldsymbol{A}$，$\boldsymbol{B}$ 是 $n\times n$ 的，$\boldsymbol{B}$ 是可逆的，$\boldsymbol{AB}$ 是可逆的. 证明 $\boldsymbol{A}$ 也是可逆的. [提示：设 $\boldsymbol{C}=\boldsymbol{AB}$，由此方程求 $\boldsymbol{A}$].
17. 假设 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 是方阵，且 $\boldsymbol{B}$ 是可逆的，由方程 $\boldsymbol{AB}=\boldsymbol{BC}$ 求 $\boldsymbol{A}$.
18. 假设 $\boldsymbol{P}$ 是可逆的，$\boldsymbol{A}=\boldsymbol{PBP}^{-1}$. 用 $\boldsymbol{A}$ 表示 $\boldsymbol{B}$.
19. 如果 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 是 $n\times n$ 可逆矩阵，方程 $\boldsymbol{C}^{-1}(\boldsymbol{A}+\boldsymbol{X})\boldsymbol{B}^{-1}=\boldsymbol{I}_n$ 是否有解 $\boldsymbol{X}$？如果有，求出该解.
20. 假设 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{X}$ 是 $n\times n$ 矩阵，$\boldsymbol{A}$，$\boldsymbol{X}$，$\boldsymbol{A}-\boldsymbol{AX}$ 可逆，假设

$$(\boldsymbol{A}-\boldsymbol{AX})^{-1}=\boldsymbol{X}^{-1}\boldsymbol{B} \tag{3}$$

   a. 解释 $\boldsymbol{B}$ 为什么是可逆的.
   b. 由(3)式求 $\boldsymbol{X}$，如果需要对矩阵取逆，解释这个矩阵为什么是可逆的.
21. 解释为什么 $n\times n$ 可逆矩阵 $\boldsymbol{A}$ 的列向量是线性无关的.
22. 解释为什么 $n\times n$ 可逆矩阵 $\boldsymbol{A}$ 的列向量可以张成 $\mathbf{R}^n$. [提示：回忆1.4节定理4]

126

23. 假设 $\boldsymbol{A}$ 是 $n\times n$ 矩阵，方程 $\boldsymbol{Ax}=\boldsymbol{0}$ 只有平凡解. 解释为什么 $\boldsymbol{A}$ 有 n 个主元列向量且 $\boldsymbol{A}$ 行等价于 $\boldsymbol{I}_n$. 根据定理7，这就表明 $\boldsymbol{A}$ 必定是可逆的.（这道习题和习题24将在2.3节中被引用）.
24. 假设 $\boldsymbol{A}$ 是 $n\times n$ 矩阵，对任意 $\boldsymbol{b}\in\mathbf{R}^n$ 方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 都有解. 解释为什么 $\boldsymbol{A}$ 一定是可逆的. [提示：$\boldsymbol{A}$ 是否行等价于 $\boldsymbol{I}_n$？]

习题25和26给出了定理4的证明，其中 $\boldsymbol{A}=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

25. 证明：如果 $ad-bc=0$，则方程 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解的个数大于1个. 这为什么表明 $\boldsymbol{A}$ 是不可逆

的？$\left[\text{提示：首先考虑 } a=b=0. \text{ 然后，如果 } a, b \text{ 不全是0，考虑向量 } \boldsymbol{x}=\begin{bmatrix} -b \\ a \end{bmatrix}.\right]$

26. 证明：如果 $ad-bc\neq 0$，则 $\boldsymbol{A}^{-1}$ 的计算公式是正确的.

对于例题5后面方框内所述的关于初等矩阵若干事实，习题27和28给出了特殊情形下的证明. 这里 $\boldsymbol{A}$ 是 3×3 矩阵，$\boldsymbol{I}=\boldsymbol{I}_3$. 一般的证明需要更多记号.

27. a. 利用2.1节方程(1)来证明：对 $i=1, 2, 3$ 有 $\text{row}_i(\boldsymbol{A})=\text{row}_i(\boldsymbol{I})\cdot\boldsymbol{A}$.
    b. 证明：如果 $\boldsymbol{A}$ 的1，2两行交换，则结果可以写成 $\boldsymbol{EA}$，其中 $\boldsymbol{E}$ 是由 $\boldsymbol{I}$ 交换1，2两行得到的初等矩阵.
    c. 证明：如果 $\boldsymbol{A}$ 的第三行乘以5，则结果可以写成 $\boldsymbol{EA}$，其中 $\boldsymbol{E}$ 是由 $\boldsymbol{I}$ 的第三行乘以5得到的矩阵.

28. 证明：如果 $\boldsymbol{A}$ 的第三行替换为 $\text{row}_3(\boldsymbol{A})-4\cdot\text{row}_1(\boldsymbol{A})$，则结果是 $\boldsymbol{EA}$，其中 $\boldsymbol{E}$ 是将 $\boldsymbol{I}$ 中 $\text{row}_3(\boldsymbol{I})$ 替换为 $\text{row}_3(\boldsymbol{I})-4\cdot\text{row}_1(\boldsymbol{I})$ 得到的矩阵.

利用本节中介绍的算法，求习题29～32中矩阵的逆(如果逆存在).

29. $\begin{bmatrix} 1 & 2 \\ 4 & 7 \end{bmatrix}$　30. $\begin{bmatrix} 5 & 10 \\ 4 & 7 \end{bmatrix}$　31. $\begin{bmatrix} 1 & 0 & -2 \\ -3 & 1 & 4 \\ 2 & -3 & 4 \end{bmatrix}$　32. $\begin{bmatrix} 1 & -2 & 1 \\ 4 & -7 & 3 \\ -2 & 6 & -4 \end{bmatrix}$

33. 利用本节的算法求 $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ 和 $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$ 的逆. 设 $\boldsymbol{A}$ 是这种形式的 $n\times n$ 矩阵，$\boldsymbol{B}$ 是它的逆. 猜想 $\boldsymbol{B}$ 的形式，然后证明 $\boldsymbol{AB}=\boldsymbol{I}$ 且 $\boldsymbol{BA}=\boldsymbol{I}$.

34. 重复习题33的方法，猜想 $\boldsymbol{A}=\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 2 & 0 & & 0 \\ 1 & 2 & 3 & & 0 \\ \vdots & & & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & n \end{bmatrix}$ 的逆，并证明你的猜测.

35. 设 $\boldsymbol{A}=\begin{bmatrix} -2 & -7 & -9 \\ 2 & 5 & 6 \\ 1 & 3 & 4 \end{bmatrix}$，只求 $\boldsymbol{A}^{-1}$ 的第三列.

36. [M]设 $\boldsymbol{A}=\begin{bmatrix} -25 & -9 & -27 \\ 546 & 180 & 537 \\ 154 & 50 & 149 \end{bmatrix}$，只求 $\boldsymbol{A}^{-1}$ 的第二、三列.

37. 设 $\boldsymbol{A}=\begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 5 \end{bmatrix}$. 构造一个 2×3 矩阵 $\boldsymbol{C}$(通过反复试验)，其中只包含元素1、-1 和0，且满足 $\boldsymbol{CA}=\boldsymbol{I}_2$. 计算 $\boldsymbol{AC}$，注意 $\boldsymbol{AC}\neq\boldsymbol{I}_3$.

38. 设 $\boldsymbol{A}=\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$. 构造一个 4×2 矩阵 $\boldsymbol{D}$，其中只包含元素1和0，且满足 $\boldsymbol{AD}=\boldsymbol{I}_2$. 是否存在一个 4×2 矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{CA}=\boldsymbol{I}_4$. 为什么？

39. 设 $\boldsymbol{D}=\begin{bmatrix}0.005 & 0.002 & 0.001\\ 0.002 & 0.004 & 0.002\\ 0.001 & 0.002 & 0.005\end{bmatrix}$是柔度矩阵，柔度单位是英寸/磅. 假设例题 3 的图 2-5 中 1，2，3 点处的作用力分别为 30，50，20 磅. 求对应各点的偏移.

40. [M]计算习题 39 中 $\boldsymbol{D}$ 的刚度矩阵 $\boldsymbol{D}^{-1}$. 列出产生下列偏移所需的作用力：点 3 的偏移为 0.04 英寸，其余点的偏移为 0.

41. [M]设 $\boldsymbol{D}=\begin{bmatrix}0.0040 & 0.0030 & 0.0010 & 0.0005\\ 0.0030 & 0.0050 & 0.0030 & 0.0010\\ 0.0010 & 0.0030 & 0.0050 & 0.0030\\ 0.0005 & 0.0010 & 0.0030 & 0.0040\end{bmatrix}$是一个弹性横梁的柔度矩阵，横梁有四个作用点. 单位是厘米/牛顿. 测量表明四个点的偏移是 0.08，0.12，0.16，0.12 厘米（见图 2-6）. 试确定四个点上的作用力.

127

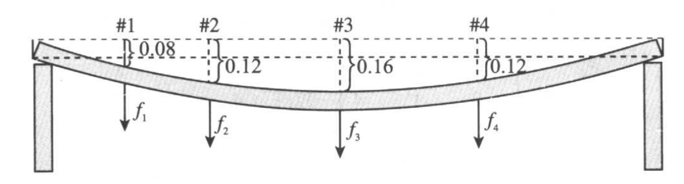


图 2-6　习题 41 和 42 中弹性横梁的偏移

42. [M]如习题 41 中的 $\boldsymbol{D}$，确定使横梁第二个点偏移 0.24 厘米，其余三点不偏移的作用力. 这个答案与 $\boldsymbol{D}^{-1}$ 中的元素有何种联系？[提示：先回答第二个点偏移 1 厘米的问题.]

基础练习答案

1. a. $\det\begin{bmatrix}3 & -9\\ 2 & 6\end{bmatrix}=3\cdot6-(-9)\cdot2=18+18=36$. 此行列式非零，所以这个矩阵是可逆的.

   b. $\det\begin{bmatrix}4 & -9\\ 0 & 5\end{bmatrix}=4\cdot5-(-9)\cdot0=20\neq0$. 这个矩阵是可逆的.

   c. $\det\begin{bmatrix}6 & -9\\ -4 & 6\end{bmatrix}=6\cdot6-(-9)(-4)=36-36=0$. 这个矩阵不可逆.

2. $[\boldsymbol{A}\quad\boldsymbol{I}]\sim\begin{bmatrix}1 & -2 & -1 & 1 & 0 & 0\\ -1 & 5 & 6 & 0 & 1 & 0\\ 5 & -4 & 5 & 0 & 0 & 1\end{bmatrix}\sim\begin{bmatrix}1 & -2 & -1 & 1 & 0 & 0\\ 0 & 3 & 5 & 1 & 1 & 0\\ 0 & 6 & 10 & -5 & 0 & 1\end{bmatrix}$

   $\sim\begin{bmatrix}1 & -2 & -1 & 1 & 0 & 0\\ 0 & 3 & 5 & 1 & 1 & 0\\ 0 & 0 & 0 & -7 & -2 & 1\end{bmatrix}$

   我们已经得到一个形如$[\boldsymbol{B}\quad\boldsymbol{D}]$的矩阵，其中 $\boldsymbol{B}$ 是含一个零行的方阵. 继续施行行变换也不可能把 $\boldsymbol{B}$ 变成 $\boldsymbol{I}$，所以我们停止变换，同时断定 $\boldsymbol{A}$ 没有逆矩阵.

2.3　可逆矩阵的性质

本节回顾了第1章中介绍的大部分概念，并且将含n个线性方程、n个未知量的

方程组与方阵建立联系．本节主要结果是定理 8.

128

【定理 8】 可逆矩阵定理

设 $\boldsymbol{A}$ 是 $n\times n$ 方阵，则下列各命题等价．即对于给定的 $\boldsymbol{A}$，下列命题或者全部为真或者全部为假.

(a) $\boldsymbol{A}$ 是可逆矩阵.

(b) $\boldsymbol{A}$ 行等价于 $n\times n$ 单位矩阵.

(c) $\boldsymbol{A}$ 有 n 个主元位置.

(d) 方程 $\boldsymbol{Ax}=\boldsymbol{0}$ 只有平凡解.

(e) $\boldsymbol{A}$ 的列向量构成一个线性无关集.

(f) 线性变换 $\boldsymbol{x}\mapsto\boldsymbol{Ax}$ 是一对一的.

(g) 对任意 $\boldsymbol{b}\in\mathbf{R}^n$ 方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 都至少有一个解.

(h) $\boldsymbol{A}$ 的列向量张成 $\mathbf{R}^n$.

(i) 线性变换 $\boldsymbol{x}\mapsto\boldsymbol{Ax}$ 将 $\mathbf{R}^n$ 映射到 $\mathbf{R}^n$ 上.

(j) 存在 $n\times n$ 矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{CA}=\boldsymbol{I}$.

(k) 存在 $n\times n$ 矩阵 $\boldsymbol{D}$ 使得 $\boldsymbol{AD}=\boldsymbol{I}$.

(l) $\boldsymbol{A}^{\mathrm{T}}$ 是可逆矩阵.

首先，我们需要一些记号．如果命题(a)为真时命题(j)总为真，我们称(a)蕴含(j)，记作(a)⇒(j)．我们将建立如图 2-7(1)所示的蕴含“圈”．如果这五个命题中的任意一个正确，则其他命题也正确．最后，我们把定理中余下的命题纳入到这个圈中.

(a) ⇒ (j) ⇒ (d) ⇒ (c) ⇒ (b) ⇒ (a)

图 2-7(1)

证明： 如果(a)是正确的，则 $\boldsymbol{A}^{-1}$ 可作为(j)中的 $\boldsymbol{C}$，所以(a)⇒(j)．另外，(j)⇒(d)可以利用 2.1 节习题 23(重新看一下这道习题)．(d)⇒(c)可以利用 2.2 节习题 23．如果 $\boldsymbol{A}$ 是方阵且有 n 个主元位置，则主元必定在主对角线上，此时 $\boldsymbol{A}$ 的简化阶梯形式是 $\boldsymbol{I}_n$．因此(c)⇒ (b)．(b)⇒(a)可以利用 2.2 节定理 7．这样就证明了图 2-7(1)中的圈.

因为 $\boldsymbol{A}^{-1}$ 可作为(k)中的 $\boldsymbol{D}$，所以(a)⇒ (k)．另外，(k)⇒(g)可以利用 2.1 节习题 24，(g)⇒(a)可以利用 2.2 节习题 24．所以(k)和(g)可以纳入这个圈中．此外，由 1.4 节定理 4 和 1.9 节定理 12(a)，(g)、(h)和(i)对任意矩阵来说都等价，因此(h)和(i)通过(g)纳入这个圈中.

(a) ⇒ (k) ⇒ (g) ⇒ (a)

(g) ⇔ (h) ⇔ (i)

(d) ⇔ (e) ⇔ (f)

(a) ⇔ (l)

图 2-7(2)

因为(d)在圈中，而(d)、(e)和(f)对任意矩阵 $\boldsymbol{A}$ 来说都等价，所以(e)和(f)也被纳入这个圈中(见 1.7 节和 1.9 节定理 12(b))．最后，(a)⇒(l)可以利用 2.2 节定理 6(c)，(l)⇒(a)可以利用同一定理(只需交换 $\boldsymbol{A}$ 和 $\boldsymbol{A}^{\mathrm{T}}$)．定理证毕. ■

根据 2.2 节定理 5，定理 8 中的命题(g)也可以写成“对任意 $\boldsymbol{b}\in\mathbf{R}^n$ 方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 都有唯一解.”这个命题自然蕴含(b)，因而蕴含：$\boldsymbol{A}$ 是可逆的.

129

下面的结论根据定理 8 和 2.2 节习题 8 得出.

设 $\boldsymbol{A}$，$\boldsymbol{B}$ 是方阵，如果 $\boldsymbol{AB}=\boldsymbol{I}$，则 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都可逆，且 $\boldsymbol{B}=\boldsymbol{A}^{-1}$，$\boldsymbol{A}=\boldsymbol{B}^{-1}$.

可逆矩阵定理把所有 $n\times n$ 矩阵的集合分成了两个不交类：可逆(非奇异)矩阵和不可逆(奇异)矩阵. 定理中的每个命题都描述了 $n\times n$ 可逆矩阵的一个性质. 定理中每个命题的否定都描述了 $n\times n$ 奇异矩阵的一个性质. 例如，一个 $n\times n$ 奇异矩阵不行等价于 $\boldsymbol{I}_n$，没有 n 个主元位置，有线性相关的列向量. 其他命题的否定将在习题中讨论.

【例题 1】　利用可逆矩阵定理判断 $\boldsymbol{A}$ 是否可逆：

$$\boldsymbol{A}=\begin{bmatrix}1 & 0 & -2\\ 3 & 1 & -2\\ -5 & -1 & 9\end{bmatrix}$$

解：

$$\boldsymbol{A}\sim\begin{bmatrix}1 & 0 & -2\\ 0 & 1 & 4\\ 0 & -1 & -1\end{bmatrix}\sim\begin{bmatrix}1 & 0 & -2\\ 0 & 1 & 4\\ 0 & 0 & 3\end{bmatrix}$$

因此 $\boldsymbol{A}$ 有三个主元位置，由可逆矩阵定理中的命题(c)，所以 $\boldsymbol{A}$ 是可逆的. ■

可逆矩阵定理的重要作用体现在它为许多重要的概念建立了联系，比如矩阵 $\boldsymbol{A}$ 中列向量的线性无关性，方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解的存在性. 但是，应该强调的是可逆矩阵定理只适用于方阵. 例如，如果一个 4×3 矩阵的列向量线性无关，我们不能利用可逆矩阵定理对方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解的存在性作任何结论.

可逆线性变换

回顾 2.1 节，矩阵乘法对应于线性变换的复合. 当 $\boldsymbol{A}$ 是可逆的，方程 $\boldsymbol{A}^{-1}\boldsymbol{Ax}=\boldsymbol{x}$ 可以看成关于线性变换的命题. 见图 2-8.

称线性变换 T：$\mathbf{R}^n\mapsto\mathbf{R}^n$ 是**可逆**(invertible)的，如果存在函数 S：$\mathbf{R}^n\mapsto\mathbf{R}^n$ 使得

$$\text{对任意 } \boldsymbol{x}\in\mathbf{R}^n,\ S(T(\boldsymbol{x}))=\boldsymbol{x} \tag{1}$$

$$\text{对任意 } \boldsymbol{x}\in\mathbf{R}^n,\ T(S(\boldsymbol{x}))=\boldsymbol{x} \tag{2}$$

下面的定理说明如果这样的 S 存在，则它是唯一的且必为线性变换. 我们称 S 为 T 的逆(inverse)，记作 T^{-1}.

130

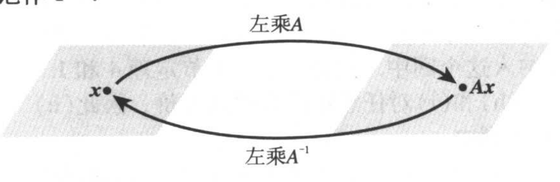


图 2-8　$\boldsymbol{A}^{-1}$ 将 $\boldsymbol{Ax}$ 变换到 $\boldsymbol{x}$

【定理 9】　设 T：$\mathbf{R}^n\mapsto\mathbf{R}^n$ 是线性变换，$\boldsymbol{A}$ 是 T 的标准矩阵. 则 T 可逆当且仅当 $\boldsymbol{A}$ 是可逆矩阵. 此时，由 $S(\boldsymbol{x})=\boldsymbol{A}^{-1}\boldsymbol{x}$ 给出的线性变换 S 是唯一满足(1)和(2)的函数.

证明：假设 T 是可逆的. 如果 $\boldsymbol{b}\in\mathbf{R}^n$ 且 $\boldsymbol{x}=S(\boldsymbol{b})$，则由(2)有 $T(\boldsymbol{x})=T(S(\boldsymbol{b}))=\boldsymbol{b}$，所以每个 $\boldsymbol{b}$ 都属于 T 的值域. 因此 T 是到 $\mathbf{R}^n$ 上的映射. 从而由可逆矩阵定理中的命

题(i)知 $\boldsymbol{A}$ 是可逆的.

反过来，假设 $\boldsymbol{A}$ 是可逆的，$S(\boldsymbol{x})=\boldsymbol{A}^{-1}\boldsymbol{x}$. 则 S 是线性变换，S 显然满足(1)和(2). 例如，

$$S(T(\boldsymbol{x}))=S(\boldsymbol{Ax})=\boldsymbol{A}^{-1}(\boldsymbol{Ax})=\boldsymbol{x}$$

因此 T 是可逆的. S 的唯一性的证明在习题 39 中有概述. ■

【例题 2】 $\mathbf{R}^n$ 到 $\mathbf{R}^n$ 的一对一线性变换 $\boldsymbol{T}$ 有什么性质?

解: T 的标准矩阵 $\boldsymbol{A}$ 各列线性无关(1.9 节定理 12). 所以 $\boldsymbol{A}$ 可逆，由可逆矩阵定理，T 是 $\mathbf{R}^n$ 到 $\mathbf{R}^n$ 上的映射. 此外，由定理 9 知 T 也是可逆的. ■

注记

实际工作中，你可能会遇到"几乎奇异"或者**病态**(ill-conditioned)的矩阵——一类特殊的可逆矩阵，当其中的元素发生细微变化后，就可能变成奇异矩阵. 这种情况下，行化简得到的主元位置可能会少于 n 个，这是舍入误差的结果. 舍入误差有时也可能会使奇异矩阵变为可逆的.

某些矩阵软件可以计算方阵的**条件数**(condition number). 条件数越大，矩阵就越近似奇异矩阵. 单位矩阵的条件数是 1，奇异矩阵的条件数是无穷. 在极端的情况下，一个矩阵软件可能区分不出奇异矩阵和病态矩阵.

习题 41 ~ 45 表明，当条件数很大时矩阵计算可能会产生相当明显的错误.

131

基础练习

1. 判断 $\boldsymbol{A}=\begin{bmatrix}2&3&4\\2&3&4\\2&3&4\end{bmatrix}$ 是否可逆.
2. 假设对某个 $n\times n$ 矩阵 $\boldsymbol{A}$，可逆矩阵定理中的命题(g)不正确. 则对 $\boldsymbol{Ax}=\boldsymbol{b}$ 形式的方程，你能得到什么结论?
3. 假设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是 $n\times n$ 矩阵，方程 $\boldsymbol{ABx}=\boldsymbol{0}$ 有一个非平凡解，则对矩阵 $\boldsymbol{AB}$，你能得到什么结论?

习题 2.3

除非特别说明，假定下面习题中的矩阵都是 $n\times n$ 矩阵. 用尽可能少的计算，判断习题 1 ~ 10 中哪些矩阵是可逆的. 并证明你的结论.

1. $\begin{bmatrix}5&7\\-3&-6\end{bmatrix}$

2. $\begin{bmatrix}-4&6\\6&-9\end{bmatrix}$

3. $\begin{bmatrix}5&0&0\\-3&-7&0\\8&5&-1\end{bmatrix}$

4. $\begin{bmatrix}-7&0&4\\3&0&-1\\2&0&9\end{bmatrix}$

5. $\begin{bmatrix}0&3&-5\\1&0&2\\-4&-9&7\end{bmatrix}$

6. $\begin{bmatrix}1&-5&-4\\0&3&4\\-3&6&0\end{bmatrix}$

7. $\begin{bmatrix}-1&-3&0&1\\3&5&8&-3\\-2&-6&3&2\\0&-1&2&1\end{bmatrix}$

8. $\begin{bmatrix}1&3&7&4\\0&5&9&6\\0&0&2&8\\0&0&0&10\end{bmatrix}$

9. [M] $\begin{bmatrix}4&0&-7&-7\\-6&1&11&9\\7&-5&10&19\\-1&2&3&-1\end{bmatrix}$

10. [M] $\begin{bmatrix}5&3&1&7&9\\6&4&2&8&-8\\7&5&3&10&9\\9&6&4&-9&-5\\8&5&2&11&4\end{bmatrix}$

习题11和12中的矩阵都是$n \times n$矩阵. 习题中的每个命题都是形如"如果<命题1>，则<命题2>"的蕴含式. 如果<命题1>为真时<命题2>总为真，我们说这个蕴含式为真. 如果存在一种情形使得<命题2>为假但<命题1>为真，则说这个蕴含式为假. 写出答案并说明理由.

11. a. 如果方程$\boldsymbol{Ax}=\boldsymbol{0}$只有平凡解. 则$\boldsymbol{A}$行等价于$n \times n$单位矩阵.
    b. 如果$\boldsymbol{A}$的列向量张成$\mathbf{R}^n$，则列向量是线性无关的.
    c. 如果$\boldsymbol{A}$是$n \times n$矩阵，则对任意$\boldsymbol{b} \in \mathbf{R}^n$方程$\boldsymbol{Ax}=\boldsymbol{b}$都至少有一个解.
    d. 如果方程$\boldsymbol{Ax}=\boldsymbol{0}$有非平凡解，则$\boldsymbol{A}$的主元位置少于$n$个.
    e. 如果$\boldsymbol{A}^{\mathrm{T}}$不可逆，则$\boldsymbol{A}$也不可逆.
12. a. 如果存在$n \times n$矩阵$\boldsymbol{D}$使得$\boldsymbol{AD}=\boldsymbol{I}$. 则存在$n \times n$矩阵$\boldsymbol{C}$使得$\boldsymbol{CA}=\boldsymbol{I}$.
    b. 如果$\boldsymbol{A}$的列向量线性无关，则$\boldsymbol{A}$的列向量张成$\mathbf{R}^n$.
    c. 如果对任意$\boldsymbol{b} \in \mathbf{R}^n$方程$\boldsymbol{Ax}=\boldsymbol{b}$都至少有一个解，则对每个向量$\boldsymbol{b}$，相应方程的解都唯一.
    d. 如果线性变换$\boldsymbol{x} \mapsto \boldsymbol{Ax}$把$\mathbf{R}^n$映到$\mathbf{R}^n$，则$\boldsymbol{A}$有$n$个主元位置.
    e. 如果存在$\boldsymbol{b} \in \mathbf{R}^n$使得方程$\boldsymbol{Ax}=\boldsymbol{b}$不相容，则线性变换$\boldsymbol{x} \mapsto \boldsymbol{Ax}$不是一对一的.
13. $m \times n$ **上三角矩阵**(upper triangular matrix)是主对角线下面的元素全为0的矩阵(如习题8). 什么时候上三角方阵可逆？证明你的结论.
14. $m \times n$ **下三角矩阵**(lower triangular matrix)是主对角线上面的元素全为0的矩阵. 什么时候下三角方阵可逆？证明你的结论.

132

15. 含有两个相同列的方阵是否可逆？为什么？
16. 当一个5×5矩阵的列向量不能张成$\mathbf{R}^5$时，它是可逆的吗？为什么？
17. 如果$\boldsymbol{A}$是可逆的，则$\boldsymbol{A}^{-1}$的列向量线性无关. 解释为什么.
18. 如果$\boldsymbol{C}$是6×6的，方程$\boldsymbol{Cx}=\boldsymbol{v}$对$\mathbf{R}^6$中的每个向量$\boldsymbol{v}$都是相容的，是否存在某个向量$\boldsymbol{v}$使得方程$\boldsymbol{Cx}=\boldsymbol{v}$的解的个数多于1个. 为什么？
19. 如果7×7矩阵$\boldsymbol{D}$的列向量线性无关，则对方程$\boldsymbol{Dx}=\boldsymbol{b}$的解会有什么样的结论？为什么？
20. 如果$n \times n$矩阵$\boldsymbol{E}$和$\boldsymbol{F}$满足$\boldsymbol{EF}=\boldsymbol{I}$. 则$\boldsymbol{E}$和$\boldsymbol{F}$交换. 解释为什么.
21. 如果存在$\mathbf{R}^n$中的某个向量$\boldsymbol{y}$使得方程$\boldsymbol{Gx}=\boldsymbol{y}$至少有一个解，$\boldsymbol{G}$的列向量能否张成$\mathbf{R}^n$，为什么？
22. 如果存在$\mathbf{R}^n$中的一个向量$\boldsymbol{c}$使得方程$\boldsymbol{Hx}=\boldsymbol{c}$不相容，则对方程$\boldsymbol{Hx}=\boldsymbol{0}$，能得到什么结论？为什么？
23. 如果$n \times n$矩阵$\boldsymbol{K}$不能行化简到$\boldsymbol{I}_n$，则对$\boldsymbol{K}$的列向量，能得到什么结论？为什么？
24. 如果$\boldsymbol{L}$是$n \times n$矩阵，方程$\boldsymbol{Lx}=\boldsymbol{0}$有平凡解，则$\boldsymbol{L}$的列向量能否张成$\mathbf{R}^n$，为什么？
25. 验证例题1前面方框内的命题.
26. 解释为什么只要$\boldsymbol{A}$的列向量线性无关，$\boldsymbol{A}^2$的列向量就可以张成$\mathbf{R}^n$.
27. 证明：如果$\boldsymbol{AB}$是可逆的，则$\boldsymbol{A}$也是可逆的. 不能利用定理6(b)，因为不能假设$\boldsymbol{A}$和$\boldsymbol{B}$是可逆的. [提示：存在矩阵$\boldsymbol{W}$使得$\boldsymbol{ABW}=\boldsymbol{I}$. 为什么?]
28. 证明：如果$\boldsymbol{AB}$是可逆的，则$\boldsymbol{B}$也是可逆的.
29. 如果$\boldsymbol{A}$是$n \times n$矩阵，且存在某个向量$\boldsymbol{b}$使得方程$\boldsymbol{Ax}=\boldsymbol{b}$至少有一个解，则线性变换$\boldsymbol{x} \mapsto \boldsymbol{Ax}$不是一对一的. 这个线性变换还有什么性质？证明你的结论.
30. 如果$\boldsymbol{A}$是$n \times n$矩阵，线性变换$\boldsymbol{x} \mapsto \boldsymbol{Ax}$是一对一的. 则这个线性变换还有什么性质？证明你的结论.
31. 假设$\boldsymbol{A}$是$n \times n$矩阵，且对$\mathbf{R}^n$中的每一个向量$\boldsymbol{b}$，方程$\boldsymbol{Ax}=\boldsymbol{b}$都至少有一个解，不利用定理5或定理8，解释为什么每个方程$\boldsymbol{Ax}=\boldsymbol{b}$实际上恰有一个解.
32. 假设$\boldsymbol{A}$是$n \times n$矩阵，方程$\boldsymbol{Ax}=\boldsymbol{0}$只有平凡解，不用可逆矩阵定理，直接解释为什么方程

$Ax=b$ 对 $\mathbf{R}^n$ 中的每一个向量 b 必定有一个解.

在习题 33 和 34 中，T 是一个 $\mathbf{R}^2$ 到 $\mathbf{R}^2$ 的线性变换. 证明 T 是可逆的，并且求出 T^{-1}.

33. $T(x_1,x_2)=(-5x_1+9x_2,4x_1-7x_2)$　　34. $T(x_1,x_2)=(6x_1-8x_2,-5x_1+7x_2)$
35. 设 T: $\mathbf{R}^n\to\mathbf{R}^n$ 是可逆线性变换，解释 T 为什么是一对一的且是满的. 利用方程(1)和(2). 然后利用一个或多个定理来作另一种解释.
36. 设 T 是把 $\mathbf{R}^n$ 映到 $\mathbf{R}^n$(满的)的线性变换，证明 T^{-1} 存在且把 $\mathbf{R}^n$ 映到 $\mathbf{R}^n$(满的)，T^{-1} 是一对一的吗?
37. 假设 T 和 U 是 $\mathbf{R}^n$ 到 $\mathbf{R}^n$ 的线性变换，使得对任意 $x\in\mathbf{R}^n$ 都有 $T(U(x))=x$，那么对任意 $x\in\mathbf{R}^n$，是否有 $U(T(x))=x$? 为什么?
38. 设线性变换 T: $\mathbf{R}^n\to\mathbf{R}^n$，存在 $\mathbf{R}^n$ 中的两个不同向量 u 和 v 使得 $T(u)=T(v)$，则 T 可能是满的吗? 为什么?
39. 设 T: $\mathbf{R}^n\to\mathbf{R}^n$ 是可逆线性变换，S 和 U 是 $\mathbf{R}^n$ 到 $\mathbf{R}^n$ 的函数，使得对任意 $x\in\mathbf{R}^n$ 都有 $S(T(x))=x$，$U(T(x))=x$. 证明：对任意 $v\in\mathbf{R}^n$ 都有 $S(v)=U(v)$. 这就表明 T 的逆是唯一的，如定理 9 所述. [提示：给定 $v\in\mathbf{R}^n$，存在某个 x 使得 $v=T(x)$. 为什么? 计算 $S(v)$ 和 $U(v)$.]
40. 假设 T 和 S 满足可逆方程(1)和(2)，其中 T 是线性变换. 直接证明 S 也是线性变换. [提示：给定 u，$v\in\mathbf{R}^n$，设 $x=S(u)$，$y=S(v)$. 则 $T(x)=u$，$T(y)=v$. 为什么? 在方程 $T(x)+T(y)=T(x+y)$ 两边同时作用 S，再考虑 $T(cx)=cT(x)$.]
41. [M]假设从实验中得到下面的方程组：

$$4.5x_1+3.1x_2=19.249 \qquad 1.6x_1+1.1x_2=6.843 \tag{3}$$

   a. 解方程组(3)，然后解下面的方程组(4)，其中等号右端的数据已四舍五入保留到小数点后两位. 求每个方程组的**精确**解.

$$4.5x_1+3.1x_2=19.25 \qquad 1.6x_1+1.1x_2=6.84 \tag{4}$$

   b. 相对于(3)中数据，(4)中对应数据的误差不超过 0.05%. 当使用(4)的解作为(3)的解的近似值时，误差百分比是多少?
   c. 利用矩阵软件求(3)的系数矩阵的条件数.

习题 42 ~ 44 展示了怎样利用矩阵 A 的条件数来估计计算方程 $Ax=b$ 的解的精确度. 如果 A 和 b 中的元素精确到 r 位有效数字，A 的条件数大约为 10^k(k 是正整数)，则计算方程 $Ax=b$ 的解时通常应该至少精确到 $r-k$ 位有效数字.

133

42. [M]求习题 9 中矩阵 A 的条件数，生成 $\mathbf{R}^4$ 中的一个随机向量 x 并且计算 $b=Ax$. 然后利用矩阵软件计算 $Ax=b$ 的解 x_1. x 和 x_1 有几位数字相同? 求出你所用的矩阵软件精确存储的数字位数，并且报告当用 x_1 代替精确值 x 时损失多少位精确数字.
43. [M]对习题 10 中的矩阵重做习题 42.
44. [M]对某个适当的 b 求解方程 $Ax=b$，以便得到五阶希尔伯特矩阵的逆的最后一列

$$A=\begin{bmatrix}1 & 1/2 & 1/3 & 1/4 & 1/5\\ 1/2 & 1/3 & 1/4 & 1/5 & 1/6\\ 1/3 & 1/4 & 1/5 & 1/6 & 1/7\\ 1/4 & 1/5 & 1/6 & 1/7 & 1/8\\ 1/5 & 1/6 & 1/7 & 1/8 & 1/9\end{bmatrix}$$

你希望在 x 的每一个元素中有多少位数字是精确的? 解释为什么. [注意：精确解是 (630, -12 600, 56 700, -88 200, 44 100.)]

45. [M]某些矩阵软件，例如 MATLAB，有一个命令可以生成不同维度的希尔伯特矩阵. 如果可能，利用一个逆命令来计算 12 阶或者更高阶希尔伯特矩阵 A 的逆矩阵. 计算 AA^{-1}，

报告你的发现.

基础练习答案

1. $\boldsymbol{A}$ 的列向量显然线性相关，因为第2列和第3列都是第1列的倍数. 因此，由可逆矩阵定理知 $\boldsymbol{A}$ 是不可逆的.
2. 如果命题(g)不正确，则 $\mathbb{R}^n$ 中至少存在一个向量 $\boldsymbol{b}$ 使得方程 $A\boldsymbol{x}=\boldsymbol{b}$ 不相容.
3. 对矩阵 $\boldsymbol{AB}$ 而不是对 $\boldsymbol{A}$ 运用可逆矩阵定理，则命题(d)就变为：方程 $AB\boldsymbol{x}=\boldsymbol{0}$ 只有平凡解，而这是不正确的. 所以 $\boldsymbol{AB}$ 是不可逆的.

2.4 分块矩阵

我们处理矩阵的一个主要特点是把矩阵 $\boldsymbol{A}$ 看成一列列向量，而不是仅仅看成一个数值阵列. 这种观点一直非常有用，现在我们希望考察 $\boldsymbol{A}$ 的其他**划分**(partition)，我们用水平分割线和竖直分割线将矩阵划分成若干块，如下面例题1所示. 分块矩阵出现在线性代数的很多现代应用中，因为这种表示方法突出了矩阵分析中的基本结构，例如本章实例介绍中的飞机设计. 本节也提供了一个复习矩阵代数和运用可逆矩阵定理的机会.

【例题1】　矩阵

134

$$A=\left[\begin{array}{ccc|cc|c} 3 & 0 & -1 & 5 & 9 & -2 \\ -5 & 2 & 4 & 0 & -3 & 1 \\ \hline -8 & -6 & 3 & 1 & 7 & -4 \end{array}\right]$$

也可以写成 2×3 **分块矩阵**(partitioned or block matrix)

$$A=\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

其中的元素是分块(或称子矩阵)：

$$A_{11}=\begin{bmatrix} 3 & 0 & -1 \\ -5 & 2 & 4 \end{bmatrix},\ A_{12}=\begin{bmatrix} 5 & 9 \\ 0 & -3 \end{bmatrix},\ A_{13}=\begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$A_{21}=[-8\quad -6\quad 3],\ A_{22}=[1\quad 7],\ A_{23}=[-4]$$ ■

【例题2】　当矩阵 $\boldsymbol{A}$ 出现在一个物理系统，比如电网，运输系统或一个大型企业的数学模型中，很自然地需要把 $\boldsymbol{A}$ 看成一个分块矩阵. 例如，如果微机电路板主要由三个VLSI(very large-scale integrated)微芯片组成，则这个电路板的矩阵可能具有如下一般形式

$$A=\left[\begin{array}{c|c|c} A_{11} & A_{12} & A_{13} \\ \hline A_{21} & A_{22} & A_{23} \\ \hline A_{31} & A_{32} & A_{33} \end{array}\right]$$

矩阵 $\boldsymbol{A}$ 的对角线上的子矩阵 A_{11}，A_{22} 和 A_{33} 是关于那三个VLSI芯片的，而其他那些子矩阵则依赖于那些微芯片之间的连接. ■

2.4.1　加法和数乘

如果矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的维度相同且划分方式也一样，则我们可以自然地对通常意义

下的矩阵和 $\boldsymbol{A}+\boldsymbol{B}$ 进行相同的划分. 此时, $\boldsymbol{A}+\boldsymbol{B}$ 的每个分块是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 中对应分块之和. 类似地, 一个数量乘以一个分块矩阵, 其运算也是逐块进行的.

2.4.2 分块矩阵的乘法

分块矩阵的乘法依然按照普通的行列法则, 就是把这些小矩阵块当作数量一样处理. 对于乘积 $\boldsymbol{AB}$, $\boldsymbol{A}$ 的列划分必须与 $\boldsymbol{B}$ 的行划分一致.

【例题3】 设

$$\boldsymbol{A}=\left[\begin{array}{ccc|cc}2 & -3 & 1 & 0 & -4\\ 1 & 5 & -2 & 3 & -1\\ \hline 0 & -4 & -2 & 7 & -1\end{array}\right]=\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix},\quad \boldsymbol{B}=\left[\begin{array}{cc}6 & 4\\ -2 & 1\\ -3 & 7\\ \hline -1 & 3\\ 5 & 2\end{array}\right]=\begin{bmatrix}\boldsymbol{B}_1\\ \boldsymbol{B}_2\end{bmatrix}$$

135

$\boldsymbol{A}$ 的 5 列被分成了两部分, 分别是 3 列和 2 列. $\boldsymbol{B}$ 的 5 行也类似地被分成两部分, 一部分 3 行, 另一部分 2 行. 我们称 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的划分**满足分块矩阵的可乘性**(conformable for block multiplication). 可以证明, 通常意义下的乘积 $\boldsymbol{AB}$ 可以写成

$$\boldsymbol{AB}=\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix}\begin{bmatrix}\boldsymbol{B}_1\\ \boldsymbol{B}_2\end{bmatrix}=\begin{bmatrix}\boldsymbol{A}_{11}\boldsymbol{B}_1+\boldsymbol{A}_{12}\boldsymbol{B}_2\\ \boldsymbol{A}_{21}\boldsymbol{B}_1+\boldsymbol{A}_{22}\boldsymbol{B}_2\end{bmatrix}=\left[\begin{array}{cc}-5 & 4\\ -6 & 2\\ \hline 2 & 1\end{array}\right]$$

由于矩阵乘法不满足交换律, 因此必须注意在乘积 $\boldsymbol{AB}$ 的表达式中, $\boldsymbol{A}$ 的子矩阵位于每个乘积项的左侧. 例如, 本题中我们有

$$\boldsymbol{A}_{11}\boldsymbol{B}_1=\begin{bmatrix}2 & -3 & 1\\ 1 & 5 & -2\end{bmatrix}\begin{bmatrix}6 & 4\\ -2 & 1\\ -3 & 7\end{bmatrix}=\begin{bmatrix}15 & 12\\ 2 & -5\end{bmatrix}$$

$$\boldsymbol{A}_{12}\boldsymbol{B}_2=\begin{bmatrix}0 & -4\\ 3 & -1\end{bmatrix}\begin{bmatrix}-1 & 3\\ 5 & 2\end{bmatrix}=\begin{bmatrix}-20 & -8\\ -8 & 7\end{bmatrix}$$

因此 $\boldsymbol{AB}$ 中最上方的分块是

$$\boldsymbol{A}_{11}\boldsymbol{B}_1+\boldsymbol{A}_{12}\boldsymbol{B}_2=\begin{bmatrix}15 & 12\\ 2 & -5\end{bmatrix}+\begin{bmatrix}-20 & -8\\ -8 & 7\end{bmatrix}=\begin{bmatrix}-5 & 4\\ -6 & 2\end{bmatrix}$$ ■

分块矩阵乘法的行列法则提供了看待两个矩阵乘积的最一般的方式. 关于矩阵乘积的下列几种观点都已经用矩阵的简单划分(即只划分行或者只划分列)进行了介绍: (1)用 $\boldsymbol{A}$ 的列定义 $\boldsymbol{Ax}$, (2)$\boldsymbol{AB}$ 的 "列定义", (3)计算 $\boldsymbol{AB}$ 的行列法则, (4)将 $\boldsymbol{AB}$ 的行看成 $\boldsymbol{A}$ 中行与矩阵 $\boldsymbol{B}$ 的乘积. 本节还将利用矩阵分块给出看待 $\boldsymbol{AB}$ 的第五种观点, 见下面的定理 10.

例题 4 中的计算为定理 10 的证明做了准备. 其中 $\mathrm{col}_k(\boldsymbol{A})$ 指 $\boldsymbol{A}$ 的第 k 列, $\mathrm{row}_k(\boldsymbol{B})$ 指 $\boldsymbol{B}$ 的第 k 行.

【例题4】 设 $\boldsymbol{A}=\begin{bmatrix}-3 & 1 & 2\\ 1 & -4 & 5\end{bmatrix}$, $\boldsymbol{B}=\begin{bmatrix}a & b\\ c & d\\ e & f\end{bmatrix}$. 证明:

$$\boldsymbol{AB}=\mathrm{col}_1(\boldsymbol{A})\mathrm{row}_1(\boldsymbol{B})+\mathrm{col}_2(\boldsymbol{A})\mathrm{row}_2(\boldsymbol{B})+\mathrm{col}_3(\boldsymbol{A})\mathrm{row}_3(\boldsymbol{B})$$

解：上面的每一项都是一个外积.（见2.1节习题27和28）由计算矩阵乘积的行列法则，

$$\text{col}_1(\boldsymbol{A})\text{row}_1(\boldsymbol{B})=\begin{bmatrix}-3\\1\end{bmatrix}[a\quad b]=\begin{bmatrix}-3a&-3b\\a&b\end{bmatrix}$$

$$\text{col}_2(\boldsymbol{A})\text{row}_2(\boldsymbol{B})=\begin{bmatrix}1\\-4\end{bmatrix}[c\quad d]=\begin{bmatrix}c&d\\-4c&-4d\end{bmatrix}$$

136

$$\text{col}_3(\boldsymbol{A})\text{row}_3(\boldsymbol{B})=\begin{bmatrix}2\\5\end{bmatrix}[e\quad f]=\begin{bmatrix}2e&2f\\5e&5f\end{bmatrix}$$

因此

$$\sum_{k=1}^{3}\text{col}_k(\boldsymbol{A})\text{row}_k(\boldsymbol{B})=\begin{bmatrix}-3a+c+2e&-3b+d+2f\\a-4c+5e&b-4d+5f\end{bmatrix}$$

这个矩阵显然就是$\boldsymbol{AB}$. 注意，$\boldsymbol{AB}$的$(1,1)$－元是上面三个外积的$(1,1)$－元之和，$\boldsymbol{AB}$的$(1,2)$－元是上面三个外积的$(1,2)$－元之和，依此类推. ■

【定理10】　$\boldsymbol{AB}$的行列展开式

如果$\boldsymbol{A}$是$m\times n$矩阵，$\boldsymbol{B}$是$n\times p$矩阵，则

$$\boldsymbol{AB}=[\text{col}_1(\boldsymbol{A})\quad\text{col}_2(\boldsymbol{A})\quad\cdots\quad\text{col}_n(\boldsymbol{A})]\begin{bmatrix}\text{row}_1(\boldsymbol{B})\\\text{row}_2(\boldsymbol{B})\\\vdots\\\text{row}_n(\boldsymbol{B})\end{bmatrix}\qquad(1)$$
$$=\text{col}_1(\boldsymbol{A})\text{row}_1(\boldsymbol{B})+\cdots+\text{col}_n(\boldsymbol{A})\text{row}_n(\boldsymbol{B})$$

证明：对每个行标i和列标j，$\text{col}_k(\boldsymbol{A})\text{row}_k(\boldsymbol{B})$的$(i,j)$－元是$\text{col}_k(\boldsymbol{A})$中$a_{ik}$和$\text{row}_k(\boldsymbol{B})$中$b_{kj}$的乘积. 因此(1)中和式的$(i,j)$－元是

$$\underset{(k=1)}{a_{i1}b_{1j}}\quad+\quad\underset{(k=2)}{a_{i2}b_{2j}}\quad+\quad\cdots\quad+\quad\underset{(k=n)}{a_{in}b_{nj}}$$

这个和式也正是由行列法则得到的乘积$\boldsymbol{AB}$的(i,j)－元. ■

2.4.3　分块矩阵的逆

下面的例题演示了分块矩阵的逆的计算过程.

【例题5】　形如

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_{11}&\boldsymbol{A}_{12}\\\boldsymbol{0}&\boldsymbol{A}_{22}\end{bmatrix}$$

的矩阵称为分块上三角矩阵. 假设$\boldsymbol{A}_{11}$是$p\times p$矩阵，$\boldsymbol{A}_{22}$是$q\times q$矩阵，$\boldsymbol{A}$是可逆的，求$\boldsymbol{A}^{-1}$.

解：将$\boldsymbol{A}^{-1}$记作$\boldsymbol{B}$，划分$\boldsymbol{B}$使得

$$\begin{bmatrix}\boldsymbol{A}_{11}&\boldsymbol{A}_{12}\\\boldsymbol{0}&\boldsymbol{A}_{22}\end{bmatrix}\begin{bmatrix}\boldsymbol{B}_{11}&\boldsymbol{B}_{12}\\\boldsymbol{B}_{21}&\boldsymbol{B}_{22}\end{bmatrix}=\begin{bmatrix}\boldsymbol{I}_p&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{I}_q\end{bmatrix}\qquad(2)$$

这个矩阵方程提供了四个方程以求解未知分块 $B_{11},\cdots,B_{22}$. 计算(2)式左端的乘积，令其元素与右端单位矩阵中的对应分块相等. 即，令 137

$$A_{11}B_{11}+A_{12}B_{21}=I_p \tag{3}$$

$$A_{11}B_{12}+A_{12}B_{22}=0 \tag{4}$$

$$A_{22}B_{21}=0 \tag{5}$$

$$A_{22}B_{22}=I_q \tag{6}$$

(6)本身并不能说明 A_{22} 是可逆的，因为我们并不知道 $B_{22}A_{22}=I_q$. 但是利用可逆矩阵定理和 A_{22} 是方阵这个事实，我们可以断定 A_{22} 是可逆的，并且 $B_{22}=A_{22}^{-1}$. 现在我们可以利用(5)得

$$B_{21}=A_{22}^{-1}0=0$$

所以(3)可以简化为

$$A_{11}B_{11}+0=I_p$$

这说明 A_{11} 是可逆的且 $B_{11}=A_{11}^{-1}$. 最后由(4)得到

$$A_{11}B_{12}=-A_{12}B_{22}=-A_{12}A_{22}^{-1}\text{ 且 }B_{12}=-A_{11}^{-1}A_{12}A_{22}^{-1}$$

因此

$$A^{-1}=\begin{bmatrix}A_{11} & A_{12}\\ 0 & A_{22}\end{bmatrix}^{-1}=\begin{bmatrix}A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1}\\ 0 & A_{22}^{-1}\end{bmatrix}$$ ■

分块对角矩阵(block diagonal matrix)是主对角线以外均为零块的分块矩阵. 分块对角矩阵是可逆的当且仅当对角线上的每个分块都可逆. 见习题13和14.

注记

1. 当矩阵太大，不适合存储在计算机内存中时，矩阵分块允许计算机每次只处理两个或三个子矩阵. 例如，一个线性规划研究小组最近为了简化一个实际问题，把矩阵划分成837行、51列. 求解这个问题在Cray超级计算机上需要花费4分钟.[1]

2. 如果允许在矩阵算法中使用分块矩阵，一些高速计算机，尤其是那些具有向量传输结构的计算机，能够更高效地运行矩阵算法.[2]

3. 高性能的数值线性代数专业软件，例如LAPACK，广泛利用了分块矩阵的运算. 138

下面的习题给出了矩阵代数的应用，并列出了应用中的典型计算.

基础练习

1. 证明 $\begin{bmatrix}I & 0\\ A & I\end{bmatrix}$ 是可逆的，并求出其逆矩阵.

---

1. 当你了解到这51个列块每个都包含大约250 000个列向量时，这个求解时间看起来就相当可观了. 原问题有837个方程和超过12 750 000多个变量！这个矩阵的逾100亿个元素中，仅有大约1亿个元素非零. 见 Robert E. Bixby et al.,"Very Large-Scale Linear Programming: A Case Study in Combining Interior Point and Simplex Methods," *Operations Research*, 40, no. 5(1992): 885~897.

2. 分块矩阵算法对计算机计算的重要性在下列文献中有介绍：*Matrix Computations*, 3rd ed., by Gene H. Golub and Charles F. van Loan (Baltimore: Johns Hopkins University Press, 1996).

2. 当 X 划分为 $[X_1 \quad X_2]$ 时，计算 X^TX.

习题 2.4

在习题1~9中，假定矩阵的划分满足分块矩阵的可乘性. 计算习题1~4中的矩阵乘积.

1. $\begin{bmatrix} I & 0 \\ E & I \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$　2. $\begin{bmatrix} E & 0 \\ 0 & F \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$　3. $\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}\begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$　4. $\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

在习题5~8中，根据 A，B，C 求 X，Y，Z 的公式，并证明你的结论. 在某些情况下，可能需要假定矩阵的维度，以得到所求的公式. [提示：计算左端的矩阵乘积，令它等于右端的矩阵.]

5. $\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ X & Y \end{bmatrix}=\begin{bmatrix} 0 & I \\ Z & 0 \end{bmatrix}$　6. $\begin{bmatrix} X & 0 \\ Y & Z \end{bmatrix}\begin{bmatrix} A & 0 \\ B & C \end{bmatrix}=\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

7. $\begin{bmatrix} X & 0 & 0 \\ Y & 0 & I \end{bmatrix}\begin{bmatrix} A & Z \\ 0 & 0 \\ B & I \end{bmatrix}=\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$　8. $\begin{bmatrix} A & B \\ 0 & I \end{bmatrix}\begin{bmatrix} X & Y & Z \\ 0 & 0 & I \end{bmatrix}=\begin{bmatrix} I & 0 & 0 \\ 0 & 0 & I \end{bmatrix}$

9. 假设 A_{11} 是可逆矩阵，求矩阵 X，Y 使下面的矩阵方程成立，并且计算 B_{22}. [提示：计算左端的矩阵乘积，令它等于右端的矩阵.]

$$\begin{bmatrix} I & 0 & 0 \\ X & I & 0 \\ Y & 0 & I \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix}=\begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \\ 0 & B_{32} \end{bmatrix}$$

10. $\begin{bmatrix} I & 0 & 0 \\ C & I & 0 \\ A & B & I \end{bmatrix}$的逆矩阵是$\begin{bmatrix} I & 0 & 0 \\ Z & I & 0 \\ X & Y & I \end{bmatrix}$，求 X，Y 和 Z.

在习题11和12中，判断下列各命题的真假，并说明理由.

11. a. 如果 $A=[A_1 \quad A_2]$ 且 $B=[B_1 \quad B_2]$，A_1，A_2 和 B_1，B_2 分别具有相同的维度，则 $A+B=[A_1+B_1 \quad A_2+B_2]$.

    b. 如果 $A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$，$B=\begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$，则 A 和 B 的划分满足分块矩阵的可乘性.

12. a. 矩阵向量乘积 Ax 的定义是分块矩阵乘法的特例.

    b. 如果 A_1，A_2，B_1 和 B_2 都是 $n\times n$ 矩阵，$A=\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$ 和 $B=[B_1 \quad B_2]$，则乘积 BA 有意义，但 AB 没有意义.

13. 设 $A=\begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}$，其中 B 和 C 是方阵. 证明：A 可逆当且仅当 B 和 C 都可逆.

14. 证明：例题5中的分块上三角矩阵 A 是可逆的当且仅当 A_{11}，A_{22} 都可逆. [提示：如果 A_{11}，A_{22} 都可逆，例题5所给的 A^{-1} 的公式就是 A 的逆矩阵.] 矩阵 A 的这个事实对于几种估计矩阵本征值的计算机算法十分重要. 本征值的概念将在第5章介绍.

15. 假设 A_{11} 是可逆的. 求 X，Y 使得

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ X & I \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} I & Y \\ 0 & I \end{bmatrix} \tag{7}$$

其中 $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$. 矩阵 S 称为 A_{11} 的舒尔补(Schur complement). 类似地，如果 A_{22} 可逆，矩阵 $A_{11} - A_{12}A_{22}^{-1}A_{21}$ 称为 A_{22} 的舒尔补. 这些表达式在系统工程等理论中经常出现. 139

16. 假设(7)式左端的分块矩阵 A 是可逆的，A_{11} 是可逆的. 证明 A_{11} 的舒尔补 S 是可逆的. [提示：(7)式右端的矩阵乘积中，外侧的两个因子总是可逆的，试加以证明.]当 A 和 A_{11} 都可逆时，利用 S^{-1}，A_{11}^{-1} 和 A 中其他元素，由(7)出发可推导出 A^{-1} 的表达式.

17. 一个太空航天探测器发射以后，可能需要调整以使探测器处在精确计算的轨道里. 雷达监测到一组向量 $x_1, \cdots, x_k$，它们给出了不同时刻探测器的实际位置与预定轨道之间偏差的信息（见图 2-9）. 令 $X_k = [x_1 \cdots x_k]$. 在雷达进行数据分析时需要计算出矩阵 $G_k = X_k X_k^T$. 一旦接收到数据向量 x_{k+1}，必须计算出新矩阵 G_{k+1}. 因为数据向量到达的速度非常快，计算的负担可能会很重. 但是分块矩阵乘法在其中发挥了巨大的作用. 计算 G_k 和 G_{k+1} 的列-行展开，说明为了更新 G_k 并得到 G_{k+1}，需要计算什么.

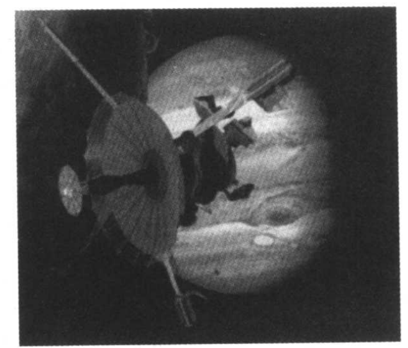
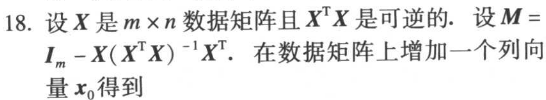


图 2-9 Galileo 探测器 1989 年 10 月 18 日发射，早在 1995 年 12 月到达木星附近



18. 设 X 是 $m \times n$ 数据矩阵且 X^TX 是可逆的. 设 $M = I_m - X(X^TX)^{-1}X^T$. 在数据矩阵上增加一个列向量 x_0 得到

$$W = [X \quad x_0]$$

计算 W^TW. 其(1,1)-元为 X^TX. 证明 X^TX 的舒尔补(习题 15)可以写成 $x_0^TMx_0$ 的形式. 可以证明，数量 $(x_0^TMx_0)^{-1}$ 是 $(W^TW)^{-1}$ 的(2,2)-元. 在适当的假设下，这个元素具有有用的统计解释.

在物理系统的工程控制研究中，一组标准的微分方程由拉普拉斯变换转化为下列线性方程组：

$$\begin{bmatrix} A - sI_n & B \\ C & I_m \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{8}$$

其中 A，B 和 C 分别是 $n \times n$，$n \times m$ 和 $m \times n$ 矩阵，s 是变量. $\mathbb{R}^m$ 中向量 u 是系统的“输入”，$\mathbb{R}^m$ 中向量 y 是“输出”，$\mathbb{R}^n$ 中向量 x 是“状态”向量(实际上，向量 x，u，y 都是 s 的函数，但是我们不考虑这个事实，因为这不影响习题 19 和 20 中的代数运算).

19. 假设 $A - sI_n$ 是可逆的，把(8)看成是含两个矩阵方程的方程组. 求解第一个方程中的 x，将其代入到第二个方程，结果是一个形如 $W(s)u = y$ 的方程，其中 $W(s)$ 是依赖于 s 的矩阵. $W(s)$ 称为这个方程组的**转移函数**，因为它把输入向量 u 转化为输出向量 y. 求解 $W(s)$，并且描述它与(8)式左端的分块**系统矩阵**有何联系. 参考习题 15.

20. 假设习题 19 中的转移函数 $W(s)$ 对某个 s 来说是可逆的. 可以证明，逆转移函数 $W(s)^{-1}$(它将输出转化为输入)是 $A - BC - sI_n$ 关于下列矩阵的舒尔补. 求这个舒尔补. 参考习题 15.

$$\begin{bmatrix} A - BC - sI_n & B \\ -C & I_m \end{bmatrix}$$

21. a. 当$A=\begin{bmatrix}1 & 0\\ 3 & -1\end{bmatrix}$，证明$A^2=I$.

b. 利用分块矩阵证明$M^2=I$. 其中$M=\begin{bmatrix}1 & 0 & 0 & 0\\ 3 & -1 & 0 & 0\\ 1 & 0 & -1 & 0\\ 0 & 1 & -3 & 1\end{bmatrix}$.

22. 推广习题21(a)[非21(b)]的思想，构造一个5×5矩阵$M=\begin{bmatrix}A & 0\\ C & D\end{bmatrix}$，使$M^2=I$且$C$为2×3非零矩阵. 证明你的构造是正确的.

23. 利用分块矩阵归纳证明：两个下三角矩阵的乘积仍然是下三角矩阵. [提示：一个$(k+1)\times(k+1)$矩阵A_1可以写成下列形式，其中a是数量，$v\in\mathbf{R}^k$，A是$k\times k$下三角矩阵. 关于归纳法的使用可参考学习指南.]

140

$$A_1=\begin{bmatrix}a & \mathbf{0}^{\mathrm{T}}\\ v & A\end{bmatrix}$$

24. 利用分块矩阵归纳证明：对$n=2,3,\cdots$，下列$n\times n$矩阵A是可逆的，B是它的逆矩阵.

$$A=\begin{bmatrix}1 & 0 & 0 & \cdots & 0\\ 1 & 1 & 0 & & 0\\ 1 & 1 & 1 & & 0\\ \vdots & & & \ddots & \\ 1 & 1 & 1 & \cdots & 1\end{bmatrix}\quad B=\begin{bmatrix}1 & 0 & 0 & \cdots & 0\\ -1 & 1 & 0 & & 0\\ 0 & -1 & 1 & & 0\\ \vdots & & \ddots & \ddots & \\ 0 & & \cdots & -1 & 1\end{bmatrix}$$

归纳时可以假设A和B是$(k+1)\times(k+1)$矩阵，然后把A和B划分为类似习题23的形式.

25. 不使用行化简，求下列矩阵的逆矩阵：

$$A=\begin{bmatrix}1 & 2 & 0 & 0 & 0\\ 3 & 5 & 0 & 0 & 0\\ 0 & 0 & 2 & 0 & 0\\ 0 & 0 & 0 & 7 & 8\\ 0 & 0 & 0 & 5 & 6\end{bmatrix}$$

26. [M]对于分块矩阵运算，可能需要输入大型矩阵的子矩阵，描述一下你所用的矩阵软件中完成下列任务的函数或命令. 假设A是一个20×30矩阵.

a. 列出A的15到20行、5到10列所构成的子矩阵.

b. 将一个5×10矩阵B插入到A中，插入的起始位置为A的第10行、第20列.

c. 构造一个形如$B=\begin{bmatrix}A & 0\\ 0 & A^{\mathrm{T}}\end{bmatrix}$的50×50矩阵. [注意：不需要详细说明$B$中的零块.]

27. [M]假设受存储器或矩阵维度的限制，你所用的矩阵软件不能处理超过32行、32列的矩阵，而某项工程涉及到50×50的矩阵A和B. 描述一下使用你所用的矩阵软件完成下列任务的命令或运算.

a. 计算$A+B$.

b. 计算AB.

c. 对$\mathbf{R}^{50}$中的某个向量b求解方程$Ax=b$，假设A可以划分成一个2×2的分块矩阵$[A_{ij}]$，其中A_{11}是20×20可逆矩阵，A_{22}是30×30可逆矩阵，A_{12}是零矩阵. [提示：给出较小的方程组来求解，不利用任何矩阵的逆.]

基础练习答案

1. 如果$\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}$是可逆的，它的逆具有形式$\begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$. 我们计算

$$\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}\begin{bmatrix} W & X \\ Y & Z \end{bmatrix} = \begin{bmatrix} W & X \\ AW+Y & AX+Z \end{bmatrix}$$

所以 W, X, Y, Z 必须满足 $W=I$, $X=0$, $AW+Y=0$ 和 $AX+Z=I$. 于是 $Y=-A$, $Z=I$. 因此

$$\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}\begin{bmatrix} I & 0 \\ -A & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

上面两个矩阵的逆序乘积也是单位矩阵，所以这个分块矩阵是可逆的，它的逆是 [141]

$\begin{bmatrix} I & 0 \\ -A & I \end{bmatrix}$. (也可以利用可逆矩阵定理.)

2. $X^{\mathrm{T}}X = \begin{bmatrix} X_1^{\mathrm{T}} \\ X_2^{\mathrm{T}} \end{bmatrix}[X_1 \quad X_2] = \begin{bmatrix} X_1^{\mathrm{T}}X_1 & X_1^{\mathrm{T}}X_2 \\ X_2^{\mathrm{T}}X_1 & X_2^{\mathrm{T}}X_2 \end{bmatrix}$. 因为 X^{T}的列就是 X 的行，所以 X^{T}和 X 的划分自动满足分块矩阵的可乘性. $X^{\mathrm{T}}X$ 的这种划分用于矩阵计算的许多计算机算法中.

2.5 矩阵分解

矩阵 A 的分解是指将 A 表示成两个或多个矩阵乘积的形式. 由于矩阵乘法涉及到数据合成(通过两个或多个线性变换的作用复合成一个矩阵)，所以矩阵分解是一种数据分解. 在计算机语言中，将 A 表示成矩阵乘积的形式相当于对 A 中的数据进行预处理，即把 A 中的数据分成两个或多个部分，使其结构在某种意义上更有用，而且更易于计算.

矩阵分解和后面介绍的线性变换分解将在本书的多个关键部分出现. 本节主要介绍的一类矩阵分解是实际应用中(例如本章实例介绍中的飞机气流问题)几种常用计算机程序的核心. 以后要学习的其他一些分解在习题中有介绍.

2.5.1 LU 分解

下面要介绍的 **LU** 分解可用于解决一类相当普遍的工商业问题——求解一系列具有相同系数矩阵的方程：

$$Ax=b_1, Ax=b_2, \cdots, Ax=b_p \tag{1}$$

习题 32 就是一个例子. 此外，5.8 节的逆幂法也是通过逐个求解形如(1)的方程，完成对矩阵本征值的估计.

当 A 可逆时，可以先计算 A^{-1}，再计算 $A^{-1}b_1$，$A^{-1}b_2$等. 然而，用行化简求解(1)中的第一个方程效率更高，同时还能得到 A 的 **LU** 分解. 然后，(1)中其余方程可用 **LU** 分解来求解.

首先，假设 A 是 $m \times n$ 矩阵，并且不使用行交换就可以行化简成阶梯形式(后面我们将处理一般情形)，则 A 可以写成 $A=LU$ 的形式，其中 L 是一个对角线元素全为 1 的 $m \times m$ 下三角矩阵，U 是 A 的 $m \times n$ 阶梯形式. 如图 2-10. 这样的分解称为 A 的 **LU 分解**(LU factorization). 矩阵 L 是可逆的，称其为单位下三角矩阵. [142]

$$A=\begin{bmatrix}1&0&0&0\\ *&1&0&0\\ *&*&1&0\\ *&*&*&1\end{bmatrix}\begin{bmatrix}\blacksquare&*&*&*&*\\ 0&\blacksquare&*&*&*\\ 0&0&0&\blacksquare&*\\ 0&0&0&0&0\end{bmatrix}$$

$$\qquad L \qquad\qquad\qquad U$$

图 2-10　LU 分解

在学习如何构造 L 和 U 之前，我们应该看一下它们为何如此重要．当 $A=LU$ 时，方程 $Ax=b$ 可以写成 $L(Ux)=b$．令 $y=Ux$，我们可以通过求解下面两个方程得到 x．

$$\boxed{\begin{array}{l}Ly=b\\ Ux=y\end{array}} \tag{2}$$

首先求解 $Ly=b$ 中的 y，然后求解 $Ux=y$ 中的 x．见图 2-11．因为 L 和 U 都是三角矩阵，所以每个方程都很容易求解．

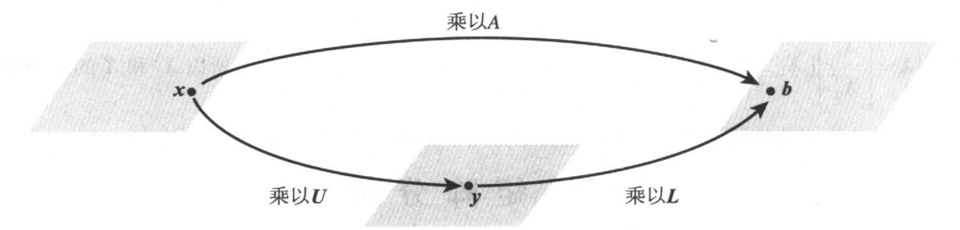


图 2-11　映射 $x\mapsto Ax$ 的分解

【例题 1】　可以证明

$$A=\begin{bmatrix}3&-7&-2&2\\ -3&5&1&0\\ 6&-4&0&-5\\ -9&5&-5&12\end{bmatrix}=\begin{bmatrix}1&0&0&0\\ -1&1&0&0\\ 2&-5&1&0\\ -3&8&3&1\end{bmatrix}\begin{bmatrix}3&-7&-2&2\\ 0&-2&-1&2\\ 0&0&-1&1\\ 0&0&0&-1\end{bmatrix}=LU$$

利用 A 的这个 LU 分解求解方程 $Ax=b$，其中 $b=\begin{bmatrix}-9\\ 5\\ 7\\ 11\end{bmatrix}$．

解： 求解 $Ly=b$ 只需在其增广矩阵的第五列上进行运算，共包括 6 次乘法和 6 次加法过程．（L 中每个主元下方的零是由所选择的行变换自动得到的．）

143

$$[L\quad b]=\begin{bmatrix}1&0&0&0&-9\\ -1&1&0&0&5\\ 2&-5&1&0&7\\ -3&8&3&1&11\end{bmatrix}\sim\begin{bmatrix}1&0&0&0&-9\\ 0&1&0&0&-4\\ 0&0&1&0&5\\ 0&0&0&1&1\end{bmatrix}=[I\quad y]$$

对于 $Ux=y$，行化简的“反向”阶段需要执行 4 次除法、6 次乘法和 6 次加法．（例如，为得到 $[U\quad y]$ 第 4 列中的零，先要对第 4 行执行 1 次除法，然后将其倍数加到上面各行，又需执行 3 次乘法、3 次加法．）

$$[U \quad y]=\begin{bmatrix}3&-7&-2&2&-9\\0&-2&-1&2&-4\\0&0&-1&1&5\\0&0&0&-1&1\end{bmatrix}\sim\begin{bmatrix}1&0&0&0&3\\0&1&0&0&4\\0&0&1&0&-6\\0&0&0&1&-1\end{bmatrix},\ x=\begin{bmatrix}3\\4\\-6\\-1\end{bmatrix}$$

不计求 L 和 U 的运算量，求 x 需要 28 次算术运算或浮点运算. 相比之下，将 $[A \quad b]$行化简为$[I \quad x]$需要 62 次运算. ■

LU 分解的计算效率取决于是否已知 L 和 U. 下面的算法表明 A 到阶梯形式 U 的行化简相当于一个 LU 分解，这是因为它能自动得到 L，不需要额外的步骤. 经过第一个行化简，L 和 U 就可用于求解系数矩阵为 A 的其他方程.

2.5.2 LU 分解算法

假设 A 可以化简成一个阶梯形式 U，化简过程仅用到行替换，即将一行的数量倍数加到这行下方的某行. 在这种情形下，存在单位下三角初等矩阵 $E_1,\cdots,E_p$使得

$$E_p\cdots E_1A=U \tag{3}$$

则

$$A=(E_p\cdots E_1)^{-1}U=LU$$

其中

$$L=(E_p\cdots E_1)^{-1} \tag{4}$$

可以证明，单位下三角矩阵的乘积和逆仍是单位下三角矩阵(见习题 19). 因此 L 是单位下三角矩阵.

注意，(3)中的行变换将 A 化简到 U，也将(4)中的 L 化简到 I，因为 $E_p\cdots E_1L=(E_p\cdots E_1)(E_p\cdots E_1)^{-1}=I$. 这个发现正是构造 L 的关键.

144

LU 分解的算法

1. 如果可能，用一系列行替换把 A 化简成阶梯形式 U.
2. 填写 L 中的元素，使得同一系列行变换序列把 L 化简到 I.

第 1 步不是总能做到的，但如果第 1 步能做到，前面的讨论说明 LU 分解存在. 例题 2 将演示第 2 步的实现过程. 根据构造，L 将满足

$$(E_p\cdots E_1)L=I$$

其中用到的初等矩阵与(3)相同，都是 $E_1,\cdots,E_p$. 因此由可逆矩阵定理，$(E_p\cdots E_1)=L^{-1}$. 由(3)，$L^{-1}A=U$，从而 $A=LU$. 所以第 2 步可以得到所需的 L.

【例题 2】 试求下列矩阵 A 的 LU 分解：

$$A=\begin{bmatrix}2&4&-1&5&-2\\-4&-5&3&-8&1\\2&-5&-4&1&8\\-6&0&7&-3&1\end{bmatrix}$$

解：因为 A 有 4 行，所以 L 应该是 4×4 矩阵. L 的第一列由 A 的第一列除以顶端的主元得到：

$$L=\begin{bmatrix}1&0&0&0\\-2&1&0&0\\1&&1&0\\-3&&&1\end{bmatrix}$$

对比A和L的第一列. 生成A第一列中的零元的那些行变换也同时生成L第一列中的零元. 我们希望这种行变换的对应对L的其他列也成立，下面观察将A化简成阶梯形式U的行变换:

$$A=\begin{bmatrix}2&4&-1&5&-2\\-4&-5&3&-8&1\\2&-5&-4&1&8\\-6&0&7&-3&1\end{bmatrix}\sim\begin{bmatrix}2&4&-1&5&-2\\0&3&1&2&-3\\0&-9&-3&-4&10\\0&12&4&12&-5\end{bmatrix}=A_1 \quad (5)$$

$$\sim A_2=\begin{bmatrix}2&4&-1&5&-2\\0&3&1&2&-3\\0&0&0&2&1\\0&0&0&4&7\end{bmatrix}\sim\begin{bmatrix}2&4&-1&5&-2\\0&3&1&2&-3\\0&0&0&2&1\\0&0&0&0&5\end{bmatrix}=U$$

上面突出标记的元素确定了A到U的行化简. 在每一个主元列中，将突出标记的元素除以这列的主元，结果记入L中:

$$\begin{bmatrix}2\\-4\\2\\-6\end{bmatrix}\begin{bmatrix}3\\-9\\12\end{bmatrix}\begin{bmatrix}2\\4\end{bmatrix}\begin{bmatrix}5\end{bmatrix}$$

$$\div 2 \quad \div 3 \quad \div 2 \quad \div 5$$

$$\downarrow \quad \downarrow \quad \downarrow \quad \downarrow$$

$$\begin{bmatrix}1&&&\\-2&1&&\\1&-3&1&\\-3&4&2&1\end{bmatrix},\ \text{于是}\ L=\begin{bmatrix}1&0&0&0\\-2&1&0&0\\1&-3&1&0\\-3&4&2&1\end{bmatrix}$$

通过简单的计算，可以验证这个L和U满足$LU=A$. ■

在实际工作中，行交换几乎总会用到，这是因为部分主元法可以保证较高的精度(回顾一下，在所有可能的主元选择中，部分主元法总是选择此列中绝对值最大的那个元素作为主元). 通过行交换，对上面的**LU**分解可作简单的修改，得到一个置换下三角矩阵L，即适当重排(置换)矩阵L的各行可以得到(单位)下三角矩阵. 利用所得的置换**LU**分解可以类似地求解方程$Ax=b$，但是$[L \quad b]$到$[I \quad y]$的化简要从L中第一列的主元开始，按照L中主元从左到右的顺序进行. 关于"**LU**分解"的文献通常都涵盖了L是置换下三角矩阵的情形. 详见学习指南.

注记

下面给出的运算量是针对$n\times n$稠密矩阵A(大多数元素非零)而言，其中n大小适中，比方说$n\geqslant 30$.[1]

1. 计算A的**LU**分解需要$2n^3/3$次浮点运算(大约与行化简$[A \quad b]$一样)，其中求A^{-1}需要$2n^3$次浮点运算.

---

1. 见 Ben Noble and James W. Daniel, *Applied Linear Algebra*, 3rd ed., (Englewood Cliffs, NJ: Prentice-Hall, 1988). 回顾一下，我们所说的浮点运算是指+，-，×或者÷.

2. 求解 $\boldsymbol{Ly}=\boldsymbol{b}$ 和 $\boldsymbol{Ux}=\boldsymbol{y}$ 大约需要 $2n^2$ 次浮点运算，这是因为任 $n\times n$ 三角方程组可以用 n^2 次浮点运算求解.

3. $\boldsymbol{b}$ 乘以 $\boldsymbol{A}^{-1}$ 也需要 $2n^2$ 次浮点运算，但结果不如从 $\boldsymbol{L}$ 和 $\boldsymbol{U}$ 得来的精确(由于计算 $\boldsymbol{A}^{-1}$ 和 $\boldsymbol{A}^{-1}\boldsymbol{b}$ 时的舍入误差).

4. 如果 $\boldsymbol{A}$ 是稀疏的(大多数元素是零)，则 $\boldsymbol{L}$ 和 $\boldsymbol{U}$ 也可能是稀疏的，而 $\boldsymbol{A}^{-1}$ 可能是稠密的. 此时，利用 LU 分解来求解方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 要比用 $\boldsymbol{A}^{-1}$ 快得多（见习题31）.

146

2.5.3 电机工程中的矩阵分解

矩阵分解与构造具有特殊性质的电网联系密切. 以下我们粗略探讨矩阵分解与电路设计间的关联.

假设图 2-12 中的方框代表某类具有输入和输出终端的电路. 用 $\begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$ 记录输入电压和输入电流(电压 v 以伏特为单位，电流 i 以安培为单位)，用 $\begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$ 记录输出电压和输出电流. 变换 $\begin{bmatrix} v_1 \\ i_1 \end{bmatrix}\mapsto\begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$ 通常是线性的. 即存在一个矩阵 A，称为*转移矩阵*，使得

$$\begin{bmatrix} v_2 \\ i_2 \end{bmatrix}=A\begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$$

图 2-13 给出了一个*梯形网络*，这是两个(也可以有多个)电路串联而成的电路，其中一个电路的输出就是另一个电路的输入. 图 2-13 左边的电路称为*串联电路*，电阻为 R_1(单位：欧姆).

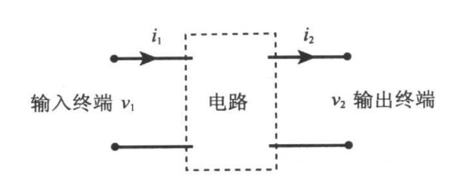


图 2-12 具有输入和输出终端的电子电路图

图 2-13 梯形网络

图 2-13 右边的电路是*并联电路*，电阻为 R_2. 利用欧姆定理和楚列斯基定律，我们可以得到串联电路和并联电路的转移矩阵，分别是

$$\underset{\text{串联电路的转移矩阵}}{\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix}} \quad \text{和} \quad \underset{\text{并联电路的转移矩阵}}{\begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}}$$

147

【例题 3】

a. 计算图 2-13 中梯形网络的转移矩阵.

b. 设计一个梯形网络，其转移矩阵是$\begin{bmatrix} 1 & -8 \\ -0.5 & 5 \end{bmatrix}$.

解：

a. 设A_1和A_2分别是串联电路和并联电路的转移矩阵. 则输入向量x先变换到A_1x，再变换到$A_2(A_1x)$. 电路的串联对应于线性变换的复合，因此梯形网络的转移矩阵是(注意顺序)

$$A_2A_1 = \begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1+R_1/R_2 \end{bmatrix} \tag{6}$$

b. 我们希望将矩阵$\begin{bmatrix} 1 & -8 \\ -0.5 & 5 \end{bmatrix}$分解成如(6)所示的若干转移矩阵的乘积. 所以我们求图2-13中的R_1和R_2，使得

$$\begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1+R_1/R_2 \end{bmatrix} = \begin{bmatrix} 1 & -8 \\ -0.5 & 5 \end{bmatrix}$$

观察上述矩阵的(1,2)－元可知$R_1=8$欧姆，观察(2,1)－元可知$1/R_2=0.5$欧姆，因此$R_2=1/0.5=2$欧姆. 代入这些值，图2-13中梯形网络就具有给定的转移矩阵. ■

电路网络的转移矩阵无需知道内部的电路结构，就能概括该网络的输入输出行为(设计规格). 为建立具有特殊性质的实体电路网络，工程师首先需要确定这样的网络是否能够构造(或实现)，然后试着将这个转移矩阵分解成若干个制造完毕等待集成的小电路的转移矩阵. 在交流电的情况下，转移矩阵的元素通常是有理复值函数(见2.4节习题19和20以及3.3节例题2). 通常的问题是寻找最小实现，即使用最少的电路元件的方案.

基础练习

求$A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix}$的LU分解. [注意：可以证明$A$只有三个主元列，所以使用例题2中的方法只能得到$L$的前三列. L的其余两列可由I_5得到.]

148

习题 2.5

在习题1～6中，利用所给的矩阵A的LU分解求解方程$Ax=b$. 习题1和2要求再用通常的行化简方法求解.

1. $A = \begin{bmatrix} 3 & -7 & -2 \\ -3 & 5 & 1 \\ 6 & -4 & 0 \end{bmatrix}$, $b = \begin{bmatrix} -7 \\ 5 \\ 2 \end{bmatrix}$　　$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -5 & 1 \end{bmatrix}\begin{bmatrix} 3 & -7 & -2 \\ 0 & -2 & -1 \\ 0 & 0 & -1 \end{bmatrix}$

2. $A=\begin{bmatrix}4&3&-5\\-4&-5&7\\8&6&-8\end{bmatrix}$, $b=\begin{bmatrix}2\\-4\\6\end{bmatrix}$　　$A=\begin{bmatrix}1&0&0\\-1&1&0\\2&0&1\end{bmatrix}\begin{bmatrix}4&3&-5\\0&-2&2\\0&0&2\end{bmatrix}$

3. $A=\begin{bmatrix}2&-1&2\\-6&0&-2\\8&-1&5\end{bmatrix}$, $b=\begin{bmatrix}1\\0\\4\end{bmatrix}$　　$A=\begin{bmatrix}1&0&0\\-3&1&0\\4&-1&1\end{bmatrix}\begin{bmatrix}2&-1&2\\0&-3&4\\0&0&1\end{bmatrix}$

4. $A=\begin{bmatrix}2&-2&4\\1&-3&1\\3&7&5\end{bmatrix}$, $b=\begin{bmatrix}0\\-5\\7\end{bmatrix}$　　$A=\begin{bmatrix}1&0&0\\1/2&1&0\\3/2&-5&1\end{bmatrix}\begin{bmatrix}2&-2&4\\0&-2&-1\\0&0&-6\end{bmatrix}$

5. $A=\begin{bmatrix}1&-2&-4&-3\\2&-7&-7&-6\\-1&2&6&4\\-4&-1&9&8\end{bmatrix}$, $b=\begin{bmatrix}1\\7\\0\\3\end{bmatrix}$　　$A=\begin{bmatrix}1&0&0&0\\2&1&0&0\\-1&0&1&0\\-4&3&-5&1\end{bmatrix}\begin{bmatrix}1&-2&-4&-3\\0&-3&1&0\\0&0&2&1\\0&0&0&1\end{bmatrix}$

6. $A=\begin{bmatrix}1&3&4&0\\-3&-6&-7&2\\3&3&0&-4\\-5&-3&2&9\end{bmatrix}$, $b=\begin{bmatrix}1\\-2\\-1\\2\end{bmatrix}$　　$A=\begin{bmatrix}1&0&0&0\\-3&1&0&0\\3&-2&1&0\\-5&4&-1&1\end{bmatrix}\begin{bmatrix}1&3&4&0\\0&3&5&2\\0&0&-2&0\\0&0&0&1\end{bmatrix}$

试求习题 7 ~ 16 中矩阵的 LU 分解(L 为单位下三角矩阵). 注意, 用 MATLAB 得到的通常都是置换 LU 分解, 这是因为为了提高数值精度, MATLAB 采用了部分主元法.

7. $\begin{bmatrix}2&5\\-3&-4\end{bmatrix}$　　8. $\begin{bmatrix}6&9\\4&5\end{bmatrix}$　　9. $\begin{bmatrix}3&-1&2\\-3&-2&10\\9&-5&6\end{bmatrix}$　　10. $\begin{bmatrix}-5&3&4\\10&-8&-9\\15&1&2\end{bmatrix}$

11. $\begin{bmatrix}3&-6&3\\6&-7&2\\-1&7&0\end{bmatrix}$　　12. $\begin{bmatrix}2&-4&2\\1&5&-4\\-6&-2&4\end{bmatrix}$　　13. $\begin{bmatrix}1&3&-5&-3\\-1&-5&8&4\\4&2&-5&-7\\-2&-4&7&5\end{bmatrix}$

14. $\begin{bmatrix}1&4&-1&5\\3&7&-2&9\\-2&-3&1&-4\\-1&6&-1&7\end{bmatrix}$　　15. $\begin{bmatrix}2&-4&4&-2\\6&-9&7&-3\\-1&-4&8&0\end{bmatrix}$　　16. $\begin{bmatrix}2&-6&6\\-4&5&-7\\3&5&-1\\-6&4&-8\\8&-3&9\end{bmatrix}$

17. 当 A 是可逆的, MATLAB 利用分解 $A=LU$ 求 A^{-1}(其中 L 可能是置换下三角矩阵), 求 L 和 U 的逆矩阵, 然后计算 $U^{-1}L^{-1}$. 利用这个方法计算习题 2 中 A 的逆矩阵(对 L 和 U 运用 2.2 节的算法).
18. 仿照习题 17 求 A^{-1}, 这里 A 是习题 3 中的矩阵.
19. 设 A 是 $n\times n$ 下三角矩阵对角线上元素非零. 证明: A 可逆且 A^{-1}是下三角矩阵. [提示: 解释为什么只用行替换和数乘就能把 A 变换到 I. (确定主元的位置)解释为什么将 A 化简到 I 的行变换也能将 I 化简到一个下三角矩阵.]
20. 设 $A=LU$ 是一个 LU 分解. 解释为什么只用行替换 A 就可以行化简到 U(这个事实是本节文中已证结果的逆).

149　21. 假设 $A=BC$，其中 B 是可逆的. 证明：任何一个将 B 化简到 I 的行变换序列也可以将 A 化简到 C. 反过来不成立，这是因为零矩阵可以分解为 $0=B\cdot 0$.

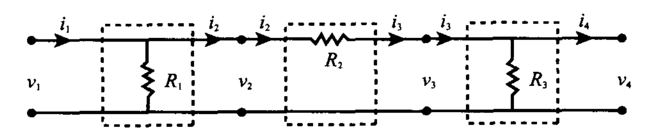
习题 22 ~ 26 给出了关于矩阵分解的一些更广泛的应用，有些在本书的后面章节还有讨论.

22. (简化 LU 分解) 对于基础练习中的矩阵 A，求一个 5×3 矩阵 B 和一个 3×4 矩阵 C 使得 $A=BC$. 试将这种思想推广到 A 是 $m\times n$ 矩阵的情形，$A=LU$，其中 U 仅有三个非零行.
23. (秩分解) 假设 $m\times n$ 矩阵 A 可以分解为 $A=CD$，其中 C 是 $m\times 4$ 的，D 是 $4\times n$ 的.
   a. 证明：A 是四个外积之和(见 2.4 节).
   b. 设 $m=400$，$n=100$. 解释为什么计算机编程者更喜欢将 A 中的数据以两个矩阵 C 和 D 的形式存储.
24. (QR 分解) 假设 $A=QR$，其中 Q 和 R 是 $n\times n$ 的，R 是可逆的上三角矩阵，Q 满足 $Q^TQ=I$. 证明：对任意 $b\in\mathbb{R}^n$，方程 $Ax=b$ 有唯一解. 利用 Q 和 R，怎样计算可以得到这个解？
25. (奇异值分解) 假设 $A=UDV^T$，其中 U 和 V 是 $n\times n$ 矩阵，且满足 $U^TU=I$ 和 $V^TV=I$，D 是对角矩阵，对角线上为正数 $\sigma_1,\cdots,\sigma_n$. 证明 A 可逆，并且给出 A^{-1} 的计算公式.
26. (谱分解) 假设 3×3 矩阵 A 可以分解为 $A=PDP^{-1}$，其中 P 是某个可逆 3×3 矩阵，D 是对角矩阵

$$D=\begin{bmatrix}1 & 0 & 0\\ 0 & 1/2 & 0\\ 0 & 0 & 1/3\end{bmatrix}$$

证明：这个分解在计算 A 的高次幂时非常有用. 利用 P 和 D 中的元素求 A^2，A^3 和 A^k (k 为正整数) 的较简洁的计算公式.

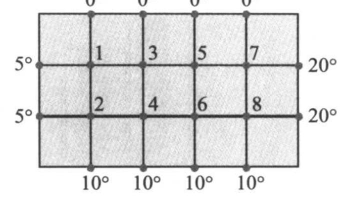
27. 设计两个不同的梯形网络，使得当输入电压为 12 伏特，输入电流为 6 安培时，输出电压均为 9 伏特，输出电流为 4 安培.
28. 证明：如果将三个并联电路(电阻为 R_1，R_2，R_3)串联起来，得到的电路网络与单个的并联电路有相同的转移矩阵. 求该电路中电阻的表达式.
29. a. 计算下图中电路网络的转移矩阵.
   b. 设 $A=\begin{bmatrix}4/3 & -12\\ -1/4 & 3\end{bmatrix}$. 利用 A 的某个适当的分解，设计一个转移矩阵为 A 的梯形网络.



30. 试求习题 29 中 A 的另一个分解，从而设计一个新的转移矩阵为 A 的梯形网络.
31. [M] 下图给出了一个金属板的稳态热传导问题，该问题的解由方程 $Ax=b$ 的解来近似，其中 $b=(5,15,0,10,0,10,20,30)$ 且

$$A=\begin{bmatrix}4 & -1 & -1 & & & & & \\ -1 & 4 & 0 & -1 & & & & \\ -1 & 0 & 4 & -1 & -1 & & & \\ & -1 & -1 & 4 & 0 & -1 & & \\ & & -1 & 0 & 4 & -1 & -1 & \\ & & & -1 & -1 & 4 & 0 & -1\\ & & & & -1 & 0 & 4 & -1\\ & & & & & -1 & -1 & 4\end{bmatrix}$$

(参考1.1节习题33)$\boldsymbol{A}$ 中未标明的元素都是零. $\boldsymbol{A}$ 中的非零元位于主对角线附近的某带状区域内. 这种带状矩阵在许多实际应用中出现, 其维度往往非常大(含数千行、列, 但非零元素都分布在相对狭窄的带状区域内).



a. 利用例题2中的方法构造 $\boldsymbol{A}$ 的LU分解, 并且注意每个因子都是带状矩阵(主对角线上方或下方有两条非零的对角线). 通过计算 $\boldsymbol{LU}-\boldsymbol{A}$ 来验证你的分解.

b. 利用这个LU分解求解方程 $\boldsymbol{Ax}=\boldsymbol{b}$.

150

c. 求 $\boldsymbol{A}^{-1}$, 注意 $\boldsymbol{A}^{-1}$ 是一个没有带状结构的稠密矩阵. 当 $\boldsymbol{A}$ 很大时, $\boldsymbol{L}$ 和 $\boldsymbol{U}$ 可以存储在比 $\boldsymbol{A}^{-1}$ 小得多的空间中. 这是我们要 $\boldsymbol{A}$ 的LU分解而不要 $\boldsymbol{A}^{-1}$ 的LU分解的另一个原因.

32. [**M**]下图中, 杆上点 $\boldsymbol{p}_1,\cdots,\boldsymbol{p}_5$ 处的温度随时间变化时, 可以用下面的带状矩阵 $\boldsymbol{A}$ 来估计杆上热量的不稳定传导.[1]

Δx Δx

p_1 p_2 p_3 p_4 p_5

矩阵中的常量 C 取决于杆的物理性能, 包括杆上相邻两点间的距离 Δx, 以及连续两次测量温度的时间间隔 Δt, 假设对于 $k=0,1,2,\cdots$, $\mathbf{R}^5$ 中的向量 $\boldsymbol{t}_k$ 列出了 $k\Delta t$ 时刻的温度. 如果杆的两端的温度保持0°, 则温度向量满足方程 $\boldsymbol{At}_{k+1}=\boldsymbol{t}_k(k=0, 1, \cdots)$, 其中

$$\boldsymbol{A}=\begin{bmatrix} (1+2C) & -C & & & \\ -C & (1+2C) & -C & & \\ & -C & (1+2C) & -C & \\ & & -C & (1+2C) & -C \\ & & & -C & (1+2C) \end{bmatrix}$$

a. 当 $C=1$ 时, 求 $\boldsymbol{A}$ 的LU分解, 形如矩阵 $\boldsymbol{A}$, 具有三条非零对角线的矩阵称为三对角矩阵. 因子 $\boldsymbol{L}$ 和 $\boldsymbol{U}$ 称为双对角矩阵.

b. 假设 $C=1$, $\boldsymbol{t}_0=(10,12,12,12,10)$. 利用 $\boldsymbol{A}$ 的这个LU分解来求温度分布 $\boldsymbol{t}_1$, $\boldsymbol{t}_2$, $\boldsymbol{t}_3$ 和 $\boldsymbol{t}_4$.

基础练习答案

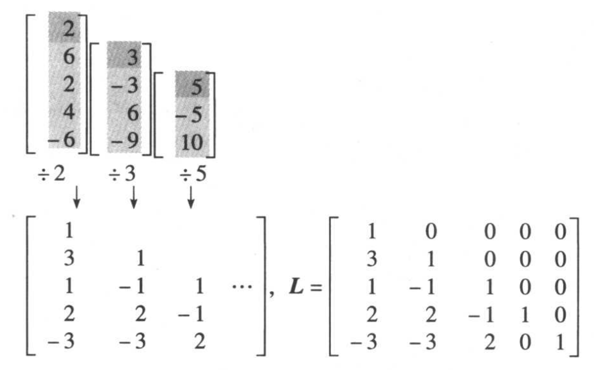
$$\boldsymbol{A}=\begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix} \sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & -3 & -1 & 6 \\ 0 & 6 & 2 & -7 \\ 0 & -9 & -3 & 13 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 10 \end{bmatrix} \sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}=\boldsymbol{U}$$

在突出标记的每一列中, 用此列的主元除以本列的每一个元素. 得到的列构成 $\boldsymbol{L}$ 的下半

1. 见 Biswa N. Datta, *Numerical Linear Algebra and Applications* (Pacific Grove, CA: Brooks/Cole, 1994), pp. 200~201.

部分的前三列. 这使得 $\boldsymbol{L}$ 到 $\boldsymbol{I}$ 的行化简对应 $\boldsymbol{A}$ 到 $\boldsymbol{U}$ 的行化简. 利用 $\boldsymbol{I}_5$ 的后两列使 $\boldsymbol{L}$ 成为单位下三角矩阵.



151

2.6 列昂惕夫投入产出模型

如第1章开头所言, 在瓦·列昂惕夫荣膺诺贝尔奖的工作中, 线性代数发挥了极其重要的作用. 本节将要介绍一种经济模型, 以此为基础建立的一些更为精细的模型正广泛应用于世界各地.

假设把国家的经济系统划分成能够生产产品或提供服务的 n 个部门, 令 $\boldsymbol{x}$ 是 $\mathbb{R}^n$ 中的**产出向量**(production vector), 它列出了各部门的年产量. 另外假设经济体系中还有一部分(称作*开放式部门*)既不生产产品也不提供服务, 而仅仅是消耗产品和服务. 令 $\boldsymbol{d}$ 是**最终需求向量**(final demand vector)或称**最终需求清单**(bill of final demands), 它列出了经济体系中非生产性部门对各部门产品和服务的需求. 向量 $\boldsymbol{d}$ 可以表示消费者需求、政府消费、生产剩余、出口以及其他外部需求.

各部门在生产产品以满足消费者需求时, 生产者对生产过程中所需投入的产品也会有**中间需求**(intermediate demand). 部门间的相互关系十分复杂, 最终需求与实际产出的联系也不明显. 列昂惕夫于是提出, 是否存在一个产出水平 $\boldsymbol{x}$, 使得产出(或者说"供应")量恰好等于产品的总需求, 这样

$$\{\text{产量 } \boldsymbol{x}\} = \{\text{中间需求}\} + \{\text{最终需求 } \boldsymbol{d}\} \tag{1}$$

列昂惕夫投入产出模型有这样一个基本假设, 那就是对每个部门都存在 $\mathbb{R}^n$ 中的一个**单位消耗向量**(unit consumption vector), 它列出了该部门每产出一个单位所需的投入. 所有的单位投入和产出都以百万美元, 而不是吨或者蒲式耳等数量来度量. 假定产品和服务的价格保持不变.

举一个简单的例子, 假设经济体系包含三个部门: 制造业、农业和服务业, 其消耗向量分别为 $\boldsymbol{c}_1$, $\boldsymbol{c}_2$ 和 $\boldsymbol{c}_3$, 如表2-1示:

表 2-1

购买自	单位产出所消耗的投入		
	制造业	农业	服务业
制造业	0.50	0.40	0.20
农业	0.20	0.30	0.10
服务业	0.10	0.10	0.30
	↑ $\boldsymbol{c}_1$	↑ $\boldsymbol{c}_2$	↑ $\boldsymbol{c}_3$

【例题 1】 制造业生产 100 单位产品需要消耗的投入量是多少？

解： 计算

$$100\boldsymbol{c}_1 = 100\begin{bmatrix}0.50\\0.20\\0.10\end{bmatrix} = \begin{bmatrix}50\\20\\10\end{bmatrix}$$

152

为生产 100 单位产品，制造业将分别从本部门其他单位、农业及服务业订购（即“需求”）并消耗 50 单位、20 单位和 10 单位. ■

若制造业决定生产 x_1 单位产品，则 $x_1\boldsymbol{c}_1$ 表示它的中间需求，因为 $x_1\boldsymbol{c}_1$ 中的量恰好是生产 x_1 单位产品的过程中的消耗量. 同样地，若以 x_2 和 x_3 记农业和服务业的计划产量，则 $x_2\boldsymbol{c}_2$ 和 $x_3\boldsymbol{c}_3$ 分别是它们的中间需求. 三个部门的中间需求总和由下式给出：

$$\{\text{中间需求}\} = x_1\boldsymbol{c}_1 + x_2\boldsymbol{c}_2 + x_3\boldsymbol{c}_3 = C\boldsymbol{x} \tag{2}$$

其中为消耗矩阵 $[\boldsymbol{c}_1 \quad \boldsymbol{c}_2 \quad \boldsymbol{c}_3]$，即

$$C = \begin{bmatrix}0.50 & 0.40 & 0.20\\0.20 & 0.30 & 0.10\\0.10 & 0.10 & 0.30\end{bmatrix} \tag{3}$$

(1)、(2)两式联立，就得到了列昂惕夫模型.

列昂惕夫投入产出模型，或列昂惕夫产出方程

$$\underset{\text{产量}}{\boldsymbol{x}} = \underset{\text{中间需求}}{C\boldsymbol{x}} + \underset{\text{最终需求}}{\boldsymbol{d}} \tag{4}$$

将 $\boldsymbol{x}$ 写成 $I\boldsymbol{x}$ 并利用矩阵代数，我们可以将(4)重写为：

$$I\boldsymbol{x} - C\boldsymbol{x} = \boldsymbol{d}$$

$$(I - C)\boldsymbol{x} = \boldsymbol{d} \tag{5}$$

【例题 2】 考虑消耗矩阵如(3)的经济体系. 假定制造业、农业和服务业的最终需求分别是 50 单位、30 单位和 20 单位. 求满足该需求的产出水平 $\boldsymbol{x}$.

解： (5)的系数矩阵是

$$I - C = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} - \begin{bmatrix}0.5 & 0.4 & 0.2\\0.2 & 0.3 & 0.1\\0.1 & 0.1 & 0.3\end{bmatrix} = \begin{bmatrix}0.5 & -0.4 & -0.2\\-0.2 & 0.7 & -0.1\\-0.1 & -0.1 & 0.7\end{bmatrix}$$

为求解(5)，对增广矩阵进行行化简：

$$\begin{bmatrix}0.5 & -0.4 & -0.2 & 50\\-0.2 & 0.7 & -0.1 & 30\\-0.1 & -0.1 & 0.7 & 20\end{bmatrix} \sim \begin{bmatrix}5 & -4 & -2 & 500\\-2 & 7 & -1 & 300\\-1 & -1 & 7 & 200\end{bmatrix} \sim \cdots \sim \begin{bmatrix}1 & 0 & 0 & 226\\0 & 1 & 0 & 119\\0 & 0 & 1 & 78\end{bmatrix}$$

最后一列四舍五入化为整数值. 制造业、农业和服务业分别需要生产约 226 单位、119 单位和 78 单位. ■

153

如果矩阵 $I-C$ 可逆，我们就可以应用 2.2 节的定理 5，其中的 A 代以 $I-C$，于是由 $(I-C)\boldsymbol{x}=\boldsymbol{d}$ 可得 $\boldsymbol{x}=(I-C)^{-1}\boldsymbol{d}$. 下面的定理表明，在大多数实际问题中，$I-C$ 可逆并且产出向量 $\boldsymbol{x}$ 是合理的，即 $\boldsymbol{x}$ 的元素非负.

定理中的概念**列和**(column sum)表示矩阵中一列元素之和. 消耗矩阵的列和一

般小于1，这是因为一个部门生产1单位产品所需的投入应该小于1单位.

【定理11】 设C是某经济体系的消耗矩阵，d是最终需求. 如果C和d的元素非负，且C的所有列和都小于1，则$(I-C)^{-1}$存在，产出向量

$$x=(I-C)^{-1}d$$

的元素非负，并且是

$$x=Cx+d$$

的唯一解.

下述讨论将解释定理为何成立，并且会给出计算$(I-C)^{-1}$的一种新方法.

2.6.1 $(I-C)^{-1}$的计算公式

设想年初时各行业就知道了需求d，并且将其产出水平设置成$x=d$，即恰好满足最终需求的水平. 当这些行业准备生产d时，他们需要订购原材料及其他投入，这就产生了中间需求Cd.

为了满足多出的这部分需求Cd，各行业又需要额外投入$C(Cd)=C^2d$. 这当然又造成了第2轮中间需求，并且当行业为满足新需求而增加产量时，又会造成第3轮需求，即$C(C^2d)=C^3d$，依此类推.

理论上，我们可以假定这个过程循环往复，不会终止，尽管在实际生活中通常不可能. 这种假想的情形可以表示为表2-2中的形式：

表 2-2

	必须满足的需求	为满足这个需求所需的投入
最终需求	d	Cd
中间需求		
第1轮需求	Cd	$C(Cd)=C^2d$
第2轮需求	C^2d	$C(C^2d)=C^3d$
第3轮需求	C^3d	$C(C^3d)=C^4d$
	⋮	⋮

154

满足全部需求的产出水平x是

$$\begin{aligned} x &= d+Cd+C^2d+C^3d+\cdots \\ &= (I+C+C^2+C^3+\cdots)d \end{aligned} \tag{6}$$

为使(6)式有意义，我们利用下列代数恒等式：

$$(I-C)(I+C+C^2+\cdots+C^m)=I-C^{m+1} \tag{7}$$

可以证明，如果C的列和都严格小于1，则$I-C$可逆，并且当m充分大时C^m近似于零矩阵，有$I-C^{m+1}\to I$. (这个事实类似于：若正数t小于1，则当m增大时有$t^m\to 0$.)利用(7)，我们可以写

$$(I-C)^{-1}\approx I+C+C^2+C^3+\cdots+C^m \tag{8}$$
当C的列和都小于1时.

(8)式意味着，只要选取足够大的m就能让等号右端充分接近$(I-C)^{-1}$.

在实际的投入产出模型中，消耗矩阵的方幂趋于零矩阵的速度相当快. 所以(8)确实给出了计算$(I-C)^{-1}$的一个有效方法. 同样地，对任意d，向量C^md也很快趋于零向量，于是(6)给出了求解$(I-C)x=d$的有效方法. 此外，如果C和d中元素非负，则(6)表明x的元素亦非负.

2.6.2 $(I-C)^{-1}$中元素在经济学上的意义

由于$(I-C)^{-1}$中的元素可以用来预测产出水平 x 随最终需求 d 的变化，因此意义重大. 事实上，$(I-C)^{-1}$第 j 列列出了对部门 j 的最终需求增加 1 个单位时各部门新增加的产量. 见习题 8.

注记

在实际应用(不仅仅是经济学)中，方程 $Ax=b$ 经常可以写成$(I-C)x=b$，其中 $C=I-A$. 如果这个方程组很大并且稀疏(元素大多为零)，C 中元素绝对值的列和可能小于 1. 这时 $C^m \to 0$. 如果 C^m趋于零的速度足够快，(6)和(8)就给出了求解 $Ax=b$ 以及求 A^{-1}的有效公式.

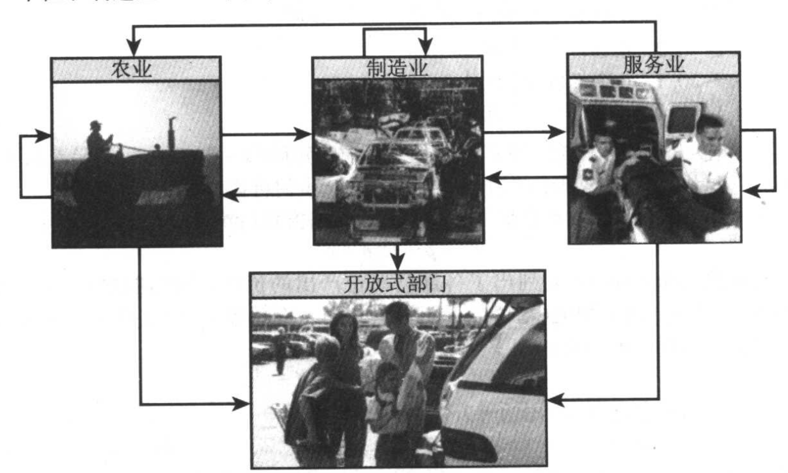
基础练习

假定一个经济体系包含两个部门：产品部门和服务部门. 产品部门的单位产出需要自身投入 0.2 单位以及服务部门投入 0.5 单位. 服务部门的单位产出需要产品部门投入 0.4 单位以及自身投入 0.3 单位. 已知最终需求是产品部门产出 20 单位，服务部门产出 30 单位. 试对这一情形建立列昂惕夫投入产出模型.

155

习题 2.6

习题 1 ~ 4 中所讨论的经济体系分为三个部门：制造业、农业和服务业. 制造业每单位产出需要消耗本部门其他企业 0.10 单位、农业 0.30 单位、服务业 0.30 单位. 农业每单位产出需要消耗自身 0.20 单位、制造业 0.60 单位、服务业 0.10 单位. 服务业每单位产出需要消耗自身 0.10 单位、制造业 0.60 单位，无需农业上的投入.



1. 建立该经济体系的消耗矩阵，确定农业产出 100 单位所产生的中间需求.
2. 若最终需求是农业产出 18 单位，其余部门不产出，试确定满足该需求的产出水平(不计算逆矩阵).
3. 若最终需求是制造业产出 18 单位，其余部门不产出，试确定满足该需求的产出水平(不计算逆矩阵).
4. 若最终需求是制造业产出 18 单位，农业产出 18 单位，服务业产出 0 单位，试确定满足该

需求的产出水平.

5. 考虑含两部门的某经济体系的产出模型 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{x}+\boldsymbol{d}$，其中

$$\boldsymbol{C}=\begin{bmatrix}0.0 & 0.5\\0.6 & 0.2\end{bmatrix},\ \boldsymbol{d}=\begin{bmatrix}50\\30\end{bmatrix}$$

利用逆矩阵确定满足最终需求的产出水平.

6. 对 $\boldsymbol{C}=\begin{bmatrix}0.1 & 0.6\\0.5 & 0.2\end{bmatrix}$，$\boldsymbol{d}=\begin{bmatrix}18\\11\end{bmatrix}$重做习题5.

7. 设 $\boldsymbol{C}$ 和 $\boldsymbol{d}$ 同习题5.

156

   a. 若最终需求是部门1产出1单位，试确定满足该需求的产出水平.

   b. 利用逆矩阵确定满足最终需求$\begin{bmatrix}51\\30\end{bmatrix}$的产出水平.

   c. 利用$\begin{bmatrix}51\\30\end{bmatrix}=\begin{bmatrix}50\\30\end{bmatrix}+\begin{bmatrix}1\\0\end{bmatrix}$这一事实，解释(a)和(b)的答案与习题5有何种关系？为什么？

8. 设 $\boldsymbol{C}$ 是各列列和均小于1的 $n\times n$ 消耗矩阵. 令 $\boldsymbol{x}$ 是满足最终需求 $\boldsymbol{d}$ 的产出向量，$\Delta\boldsymbol{x}$ 是满足另一最终需求 $\Delta\boldsymbol{d}$ 的产出向量.

   a. 证明，若最终需求由变 $\boldsymbol{d}$ 至 $\boldsymbol{d}+\Delta\boldsymbol{d}$，则新的产出水平是 $\boldsymbol{x}+\Delta\boldsymbol{x}$. 这样，$\Delta\boldsymbol{x}$ 给出了需求变动引发的产出变动.

   b. 令 $\Delta\boldsymbol{d}$ 是 $\mathbf{R}^n$ 中首元素为1，其余元素均为0的向量. 解释其对应的产出 $\Delta\boldsymbol{x}$ 为什么是 $(\boldsymbol{I}-\boldsymbol{C})^{-1}$的首列. 这表明$(\boldsymbol{I}-\boldsymbol{C})^{-1}$的首列列出了对部门1的最终需求增加1个单位时各部门新增加的产量.

9. 求解含三部门的某经济体系的列昂惕夫产出方程，已知

$$\boldsymbol{C}=\begin{bmatrix}0.2 & 0.2 & 0.0\\0.3 & 0.1 & 0.3\\0.1 & 0.0 & 0.2\end{bmatrix}\text{和 }\boldsymbol{d}=\begin{bmatrix}40\\60\\80\end{bmatrix}$$

10. 1972年美国经济体系的消耗矩阵 $\boldsymbol{C}$ 满足如下性质：矩阵$(\boldsymbol{I}-\boldsymbol{C})^{-1}$的每个元素都非零(并且是正的).[1] 这说明提高对单个部门产出的需求会产生何种影响?

11. 列昂惕夫产出方程 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{x}+\boldsymbol{d}$ 通常与下列对偶的**价格方程**(price equation)相伴：

$$\boldsymbol{p}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{p}+\boldsymbol{v}$$

其中**价格向量**(price vector)$\boldsymbol{p}$ 列出了各部门单位产出的价格，**增值向量**(value added vector)$\boldsymbol{v}$ 列出了单位产出的增值(增值包括工资、利润、折旧等). 经济学中一个重要的事实是，国内生产总值(GDP)可以用下列两种方式表示：

$$\{\text{国内生产总值}\}=\boldsymbol{p}^{\mathrm{T}}\boldsymbol{d}=\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x}$$

证明第二个等号成立. [提示:用两种方法计算 $\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}$.]

12. 设 $\boldsymbol{C}$ 为一个消耗矩阵,且当 $m\to\infty$ 时有 $\boldsymbol{C}^m\to 0$. 对 $m=1,2,\cdots$，令 $\boldsymbol{D}_m=\boldsymbol{I}+\boldsymbol{C}+\cdots+\boldsymbol{C}^m$. 求关联 $\boldsymbol{D}_m$和 $\boldsymbol{D}_{m+1}$的一个差分方程，进而得到利用公式(8)计算$(\boldsymbol{I}-\boldsymbol{C})^{-1}$的一个迭代过程.

13. [M] 下面的消耗矩阵 $\boldsymbol{C}$ 建立在1958年美国经济投入产出数据之上，81个部门的数据被划分成7大类：(1)非金属日用品和个人产品，(2)高级金属制品(比如机动车)，(3)初级

1. Wassily W. Leontief, "The World Economy of the Year 2000", *Scientific American*, September 1980, pp. 206~231.

金属制品和矿业，(4)初级非金属制品和农业，(5)能源，(6)服务业以及(7)娱乐业和综合性产业.[1] 试求满足最终需求 d 的产出水平(单位均为百万美元).

$$\begin{bmatrix} 0.1588 & 0.0064 & 0.0025 & 0.0304 & 0.0014 & 0.0083 & 0.1594 \\ 0.0057 & 0.2645 & 0.0436 & 0.0099 & 0.0083 & 0.0201 & 0.3413 \\ 0.0264 & 0.1506 & 0.3557 & 0.0139 & 0.0142 & 0.0070 & 0.0236 \\ 0.3299 & 0.0565 & 0.0495 & 0.3636 & 0.0204 & 0.0483 & 0.0649 \\ 0.0089 & 0.0081 & 0.0333 & 0.0295 & 0.3412 & 0.0237 & 0.0020 \\ 0.1190 & 0.0901 & 0.0996 & 0.1260 & 0.1722 & 0.2368 & 0.3369 \\ 0.0063 & 0.0126 & 0.0196 & 0.0098 & 0.0064 & 0.0132 & 0.0012 \end{bmatrix}, \boldsymbol{d} = \begin{bmatrix} 74\,000 \\ 56\,000 \\ 10\,500 \\ 25\,000 \\ 17\,500 \\ 196\,000 \\ 5\,000 \end{bmatrix}$$

14. [M]习题 13 中的需求向量对于 1958 年的经济数据来说是合理的，不过，在所引用的文献中，列昂惕夫用于讨论的需求向量更接近 1964 年的数据：

$$\boldsymbol{d} = (99640, 75548, 14444, 33501, 23527, 263985, 6526)$$

试求满足这一需求的产出水平.

15. [M]利用(6)式解习题 13. 令 $\boldsymbol{x}^{(0)} = \boldsymbol{d}$，对 $k = 1, 2, \cdots$，计算 $\boldsymbol{x}^{(k)} = \boldsymbol{d} + \boldsymbol{C}\boldsymbol{x}^{(k-1)}$. 经过多少步可以得到习题 13 保留四位有效数字的结果？

157

基础练习答案

已知表 2-3 的数据：

表 2-3

购买自	单位产出所需的投入		额外需求
	产品部门	服务部门	
产品部门	0.2	0.4	20
服务部门	0.5	0.3	30

列昂惕夫投入产出模型是 $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{x} + \boldsymbol{d}$，其中

$$\boldsymbol{C} = \begin{bmatrix} 0.2 & 0.4 \\ 0.5 & 0.3 \end{bmatrix}, \quad \boldsymbol{d} = \begin{bmatrix} 20 \\ 30 \end{bmatrix}$$

2.7 计算机图形学上的应用

计算机图形即计算机屏幕上显示或放映的图像. 计算机图形学的应用相当普遍，并且发展迅速. 例如，计算机辅助设计(CAD)已成为许多工业流程的主体部分，比如本章实例介绍中提到的飞机设计过程. 计算机图形在娱乐业的使用十分可观，从黑客帝国中的特技到游戏机 PlayStation 2 和 Xbox，无一不使用了这项技术.

工商业中，大多数交互式计算机软件在屏幕显示、数据的图形显示、桌面印刷系统以及商业和教育类幻灯片制作等方面都使用了计算机图形. 因此，任何学习计算机语言的人都应该学习计算机图形学，并且至少掌握二维(2D)图形的使用.

本节将考察在处理和显示计算机图形时所需的数学基础知识，比如飞机的曲线框架模型. 这样一幅图像(或图片)包含大量的点、连接线(直线或曲线)以及连接线所围成封闭区域的填充方式信息. 曲线往往用一些短的直线段来近似，图则被定义成数学上的一个点列.

---

1. Wassily W. Leontief, "The Structure of the U. S. Economy", *Scientific American*, April 1965, pp. 30 ~ 32.

屏幕上呈现的字母是最简单的 2D 图形符号．它们中某些被存储成曲线框架对象；对于包含弯曲部分的字母还另外存储了曲线的数学公式．

【例题 1】　图 2-14 中的大写字母 N 由 8 个点或顶点确定．点的坐标可以存储在数值矩阵 $\boldsymbol{D}$ 中．

$$\begin{array}{r} \\ x\text{ 坐标} \\ y\text{ 坐标} \end{array} \begin{array}{c} \text{顶点：} \\ \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{bmatrix} 0 & 0.5 & 0.5 & 6 & 6 & 5.5 & 5.5 & 0 \\ 0 & 0 & 6.42 & 0 & 8 & 8 & 1.58 & 8 \end{bmatrix} = \boldsymbol{D} \end{array}$$

158

实际上，除了 $\boldsymbol{D}$ 以外，还需要详细说明哪些点之间有直线相连，此处略过． ■

图 2-14　正体的 N

使用直线段描述图形对象的主要原因是，计算机图形的标准变换将直线段映射为直线段(例如可参考 1.8 节习题 27)．一个对象只要对其顶点施以变换，然后用适当的直线连结各顶点的像，就能得到这个对象在该变换下完整的像．

【例题 2】　已知 $\boldsymbol{A} = \begin{bmatrix} 1 & 0.25 \\ 0 & 1 \end{bmatrix}$，描述剪切变换 $\boldsymbol{x} \mapsto \boldsymbol{Ax}$ 在例题 1 中字母 N 上的作用．

解：根据矩阵乘法的定义，乘积 $\boldsymbol{AD}$ 的列就是字母 N 各顶点的像．

$$\begin{array}{c} \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \boldsymbol{AD} = \begin{bmatrix} 0 & 0.5 & 2.105 & 6 & 8 & 7.5 & 5.895 & 2 \\ 0 & 0 & 6.420 & 0 & 8 & 8 & 1.580 & 8 \end{bmatrix} \end{array}$$

图 2-15 中绘出了变换后的顶点以及与原图对应的连接线． ■

图 2-15 中斜体的 N 看起来太宽了，我们可以用压缩变换压缩其宽度。

图 2-15　斜体的 N

【例题 3】　求一个矩阵变换，它先施行例题 2 中的剪切变换，然后再将全部 x 坐标乘以 0.75．

解：将一个点的 x 坐标乘以 0.75，这个变换对应的矩阵是

$$\boldsymbol{S} = \begin{bmatrix} 0.75 & 0 \\ 0 & 1 \end{bmatrix}$$

于是复合变换为

$$\boldsymbol{SA} = \begin{bmatrix} 0.75 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0.25 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.75 & 0.1875 \\ 0 & 1 \end{bmatrix}$$

复合变换作用的结果如图 2-16 所示． ■

图 2-16　N 的复合变换

计算机图形所用的数学主要是矩阵乘法．但是，屏幕上对象的平移并不能直接对应于矩阵乘法，因为平移不是线性变换．消除这一障碍的标准做法是引入所谓的齐次坐标．

2.7.1　齐次坐标

159

$\mathbf{R}^2$ 中每个点 (x,y) 都可以等同于点 $(x,y,1)$，后者是 $\mathbf{R}^2$

中 xy 平面上方 1 单位的平面上的点. 我们称(x,y)有齐次坐标$(x,y,1)$. 例如，点$(0,0)$有齐次坐标$(0,0,1)$. 点的齐次坐标不能进行加法和数乘运算，但可以通过 3×3矩阵进行变换.

【例题 4】 形如$(x,y)\mapsto(x+h,\ y+k)$的平移用齐次坐标可以写成$(x,y,1)\mapsto(x+h,\ y+k,1)$(见图 2-17). 这一变换可用下列矩阵乘法表示：

$$\begin{bmatrix} 1 & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}=\begin{bmatrix} x+h \\ y+k \\ 1 \end{bmatrix}$$ ■

【例题 5】 利用齐次坐标，$\mathbf{R}^2$中的任意线性变换都可以表示为形如$\begin{bmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}$的矩阵，其中 $\boldsymbol{A}$ 是 2×2 矩阵. 下面是一些典型的例子：

$$\begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix},\quad \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},\quad \begin{bmatrix} s & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

绕原点逆时针旋转角度 φ　　关于 $y=x$ 轴对称　　x 坐标数乘 s，y 坐标数乘 t

■

2.7.2 复合变换

计算机屏幕上图形的移动往往需要经过两次或多次变换. 当我们采用齐次坐标时，这些变换的合成对应于矩阵乘法.

【例题 6】 求一个 3×3 矩阵，它对应于下列变换的复合变换：数乘 0.3，旋转 90°，最后将图形上每个点的坐标加上$(-0.5,2)$. 见图 2-18.

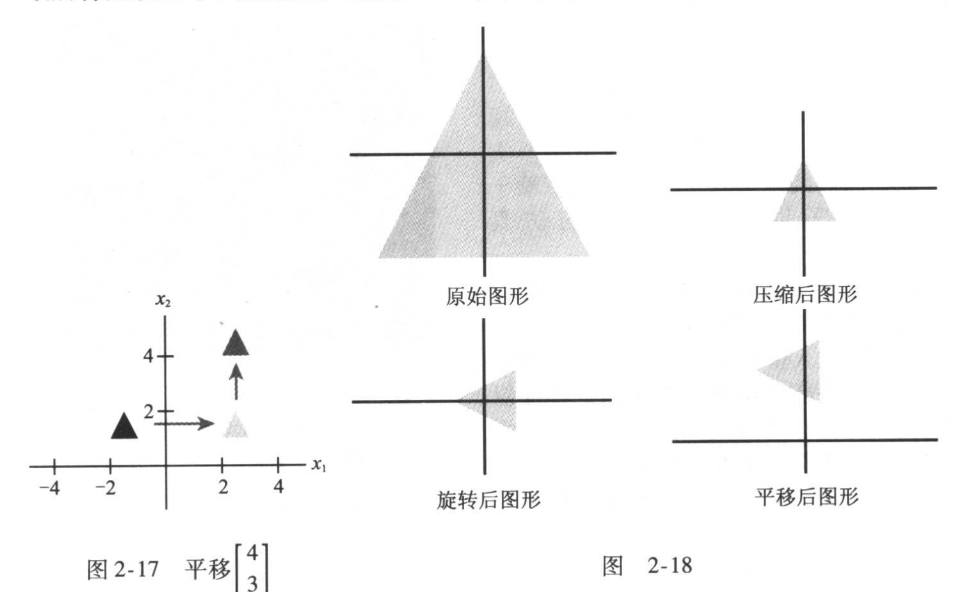


图 2-17 平移$\begin{bmatrix} 4 \\ 3 \end{bmatrix}$　　　图 2-18

解：若 $\varphi=\pi/2$，则 $\sin\varphi=1$ 且 $\cos\varphi=0$. 根据例题4和5，我们有

$$\begin{bmatrix}x\\y\\1\end{bmatrix}\xrightarrow{\text{数乘}}\begin{bmatrix}0.3&0&0\\0&0.3&0\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\1\end{bmatrix}$$

$$\xrightarrow{\text{旋转}}\begin{bmatrix}0&-1&0\\1&0&0\\0&0&1\end{bmatrix}\begin{bmatrix}0.3&0&0\\0&0.3&0\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\1\end{bmatrix}$$

$$\xrightarrow{\text{平移}}\begin{bmatrix}1&0&-0.5\\0&1&2\\0&0&1\end{bmatrix}\begin{bmatrix}0&-1&0\\1&0&0\\0&0&1\end{bmatrix}\begin{bmatrix}0.3&0&0\\0&0.3&0\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\1\end{bmatrix}$$

160

复合变换对应的矩阵是

$$\begin{bmatrix}1&0&-0.5\\0&1&2\\0&0&1\end{bmatrix}\begin{bmatrix}0&-1&0\\1&0&0\\0&0&1\end{bmatrix}\begin{bmatrix}0.3&0&0\\0&0.3&0\\0&0&1\end{bmatrix}$$

$$=\begin{bmatrix}0&-1&-0.5\\1&0&2\\0&0&1\end{bmatrix}\begin{bmatrix}0.3&0&0\\0&0.3&0\\0&0&1\end{bmatrix}=\begin{bmatrix}0&-0.3&-0.5\\0.3&0&2\\0&0&1\end{bmatrix}$$

■

2.7.3　3D计算机图形

计算机图形学中一些最新最有趣的工作与分子建模相关. 生物学家利用3D(三维)图形来研究仿真蛋白质分子，并且在这些分子上寻找可以容纳药物分子的活性部位. 他们通过旋转和平移实验药物，使之附着在蛋白质上. 蛋白质化学反应的这种可视化技术对现代医药和癌症研究意义重大. 实际上，药物制造的改进在一定程度上依赖于计算机对分子以及交互作用的仿真能力.[1]

图2-19　虚拟现实中的分子建模(北卡罗莱纳大学Chape Hill分校计算机科学系. Bo Strain摄)

当前分子建模的研究集中于虚拟现实，在虚拟现实的环境中研究者可以看见并且感觉到药物分子融入蛋白质. 图2-19中的触觉反馈就是由一种能将压力可视化的远程操纵器所提供. 另一种虚拟现实提供了头盔和手套，用以监测头部、手和手指的运动. 头盔上每只眼睛的部位都有一个小型计算机屏幕. 如何让虚拟环境更加真实？这是对工

161

1. Robert Pool, "Computing in Science", *Science* **256**, 3 April 1992, p. 45.

程设计师、科学家和数学家们共同的挑战. 这里我们所考察的数学知识仅仅打开了一扇通往这一研究领域的大门.

2.7.4 齐次 3D 坐标

与 2D 的情况类似，我们称$(x,y,z,1)$是$\mathbf{R}^3$中点(x,y,z)的齐次坐标. 一般地，称(X,Y,Z,H)是(x,y,z)的**齐次坐标**(homogeneous coordinates)，如果$H\neq 0$且

$$x=\frac{X}{H},\qquad y=\frac{Y}{H},\ 和\ z=\frac{Z}{H} \tag{1}$$

$(x,y,z,1)$的任意非零数量倍都给出了(x,y,z)的一套齐次坐标. 例如，$(10,-6,14,1)$和$(-15,9,-21,-3)$都是$(5,-3,7)$的齐次坐标.

下面的例题描述了分子建模过程中将药物分子移动到蛋白质分子内的变换.

【例题 7】 给出下列变换对应的 4×4 矩阵.

a. 绕 y 轴旋转 30°(根据习惯，正的角度值代表沿旋转轴(此时是 y 轴)正半轴俯视原点时的逆时针方向).

b. 平移向量 $\boldsymbol{p}=(-6,4,5)$.

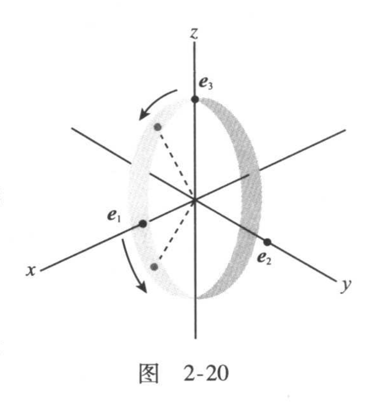


图 2-20

解:

a. 首先构造这个旋转的 3×3 矩阵. 向量 $\boldsymbol{e}_1$ 向 z 负轴方向旋转，直至$(\cos 30°,0,-\sin 30°)=(\sqrt{3}/2,0,-0.5)$. y 轴上的向量 $\boldsymbol{e}_1$ 不移动，而 z 轴上的 $\boldsymbol{e}_3$ 向 x 正轴方向旋转，直至$(\sin 30°,0,\cos 30°)=(0.5,0,\sqrt{3}/2)$. 见图 2-20. 根据 1.9 节可知，这个旋转的标准矩阵是

$$\boldsymbol{A}=\begin{bmatrix}\sqrt{3}/2 & 0 & 0.5\\ 0 & 1 & 0\\ -0.5 & 0 & \sqrt{3}/2\end{bmatrix}$$

于是齐次坐标的旋转矩阵是

$$\boldsymbol{A}=\begin{bmatrix}\sqrt{3}/2 & 0 & 0.5 & 0\\ 0 & 1 & 0 & 0\\ -0.5 & 0 & \sqrt{3}/2 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

162

b. 我们希望点$(x,y,z,1)$移至$(x-6,\ y+4,\ z+5,\ 1)$. 这个变换所对应的矩阵是

$$\begin{bmatrix}1 & 0 & 0 & -6\\ 0 & 1 & 0 & 4\\ 0 & 0 & 1 & 5\\ 0 & 0 & 0 & 1\end{bmatrix}$$

■

2.7.5 透视投影

在二维的计算机屏幕上呈现一个三维物体时，需要将物体投影在某个视平面上

(这里我们略过其他一些重要的步骤，比如选择呈现在屏幕上的视平面比例). 为简单起见，以 xy 平面代表计算机屏幕，并假定观察者的眼睛位于 z 正轴上的点 $(0,0,d)$. 透视投影将点 (x,y,z) 映射到像点 $(x^*,y^*,0)$ 上，这两个点与眼睛的位置共线，其中眼睛的位置又被称作投影中心. 如图 2-21a.

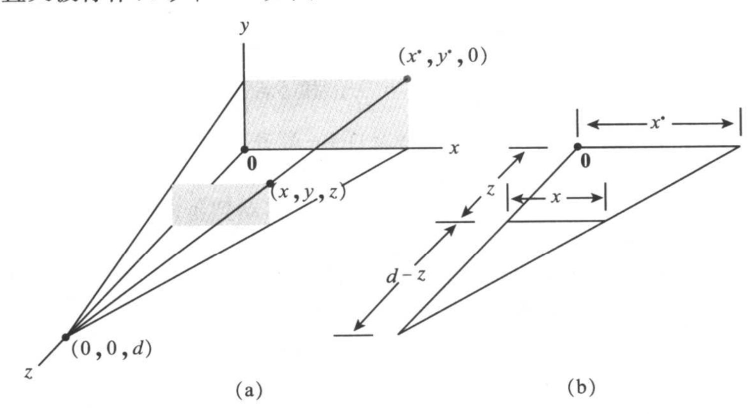


图 2-21　(x,y,z) 到 $(x^*,y^*,0)$ 上的透视投影

图 2-21b 重绘了 a 中 xy 平面上的三角形，并且标记了各线段长度. 两三角形相似表明

$$\frac{x^*}{d}=\frac{x}{d-z}\text{和 } x^*=\frac{dx}{d-z}=\frac{x}{1-z/d}$$

同样还有

163

$$y^*=\frac{y}{1-z/d}$$

利用齐次坐标，我们可以将透视投影表示成矩阵 $\boldsymbol{P}$. 我们希望将 $(x,y,z,1)$ 映射到 $\left(\frac{x}{1-z/d},\ \frac{y}{1-z/d},\ 0,\ 1\right)$. 用数 $1-z/d$ 乘以这些坐标，则可用 $(x,y,0,1-z/d)$ 作为像的齐次坐标. 现在很容易表示出 $\boldsymbol{P}$，实际上，

$$\boldsymbol{P}\begin{bmatrix}x\\y\\z\\1\end{bmatrix}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&-1/d&1\end{bmatrix}\begin{bmatrix}x\\y\\z\\1\end{bmatrix}=\begin{bmatrix}x\\y\\0\\1-z/d\end{bmatrix}$$

【例题 8】　设 S 是顶点为 $(3,1,5)$, $(5,1,5)$, $(5,0,5)$, $(3,0,5)$, $(3,1,4)$, $(5,1,4)$, $(5,0,4)$ 和 $(3,0,4)$ 的盒子. 求以 $(0,0,10)$ 为投影中心时，S 在透视投影下的像.（见图 2-22）

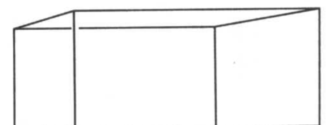


图 2-22　透视投影下的 S

解: 令 $\boldsymbol{P}$ 是投影矩阵，$\boldsymbol{D}$ 是齐次坐标表示下的 S 的数据矩阵. 则 S 投影后像的数据矩阵是

$$\boldsymbol{PD}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&-1/10&1\end{bmatrix}\overset{\text{顶点}\quad \mathbf{1}\ \mathbf{2}\ \mathbf{3}\ \mathbf{4}\ \mathbf{5}\ \mathbf{6}\ \mathbf{7}\ \mathbf{8}}{\begin{bmatrix}3&5&5&3&3&5&5&3\\1&1&0&0&1&1&0&0\\5&5&5&5&4&4&4&4\\1&1&1&1&1&1&1&1\end{bmatrix}}$$

$$=\begin{bmatrix}3&5&5&3&3&5&5&3\\1&1&0&0&1&1&0&0\\0&0&0&0&0&0&0&0\\0.5&0.5&0.5&0.5&0.6&0.6&0.6&0.6\end{bmatrix}$$

利用(1)，将每列头三个元素除以该列第四行上的元素，就可以得到投影像在 $\mathbf{R}^3$ 下的坐标：

$$\overset{\text{顶点}\quad \mathbf{1}\ \mathbf{2}\ \mathbf{3}\ \mathbf{4}\ \mathbf{5}\ \mathbf{6}\ \mathbf{7}\ \mathbf{8}}{\begin{bmatrix}6&10&10&6&5&8.3&8.3&5\\2&2&0&0&1.7&1.7&0&0\\0&0&0&0&0&0&0&0\end{bmatrix}}$$

■

本书的网站上给出了计算机图形学的一些有趣应用，也对透视投影进行了进一步讨论. 网站上的一个计算机程序还包含了简单的动画.

164

注记

计算机图形学中，3D 物体的连续运动涉及大量 4×4 矩阵计算，尤其是要求物体表面看起来真实、有纹理且光线自然的时候，运算量更大. 高端的计算机图形平台在微芯片和电路中嵌入了 4×4 矩阵运算和图形算法. 它们每秒钟可以执行数十亿次矩阵乘法，可用于实现 3D 游戏软件中的仿真彩色动画.[1]

更多阅读

James D. Foley, Andries van Dam, Steven K. Feiner, and John F. Hughes, *Computer Graphics: Principles and Practice*, 3rd ed. (Boston, MA: Addison-Wesley, 2002), Chapters 5 and 6.

基础练习

$\mathbf{R}^2$ 中图形绕点 $\boldsymbol{p}$ 的旋转可经由下列变换实现：平移向量 $-\boldsymbol{p}$，绕原点旋转，平移向量 $\boldsymbol{p}$. 见图 2-23. 利用齐次坐标，建立将点绕$(-2,6)$旋转 $-30°$的 3×3 矩阵.

---

1. 参考 Jan Ozer, "High-Performance Graphics Boards", *PC Magazine* **19**, 1 September 2000, pp. 187 ~ 200 以及 "The Ultimate Upgrade Guide: Moving On Up", *PC Magazine* **21**, 29 January 2002, pp. 82 ~ 91.

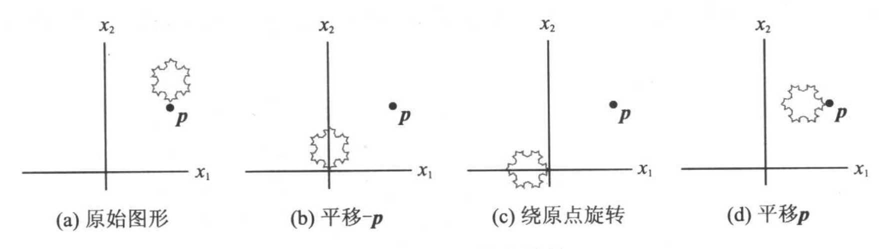


图 2-23 图形绕点旋转

习题 2.7

1. 哪一个 3×3 矩阵在 $\mathbf{R}^2$ 齐次坐标上的作用与例题 2 中剪切矩阵 A 的作用相同？
2. 已知三角形数据矩阵为 $\boldsymbol{D}=\begin{bmatrix}5&2&4\\0&2&3\end{bmatrix}$，利用矩阵乘法求它在关于 y 轴的对称下的像. 绘出原始三角形及其像的草图.

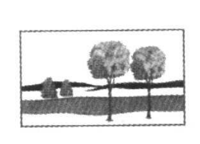
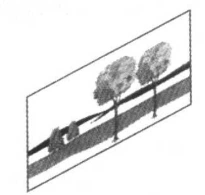
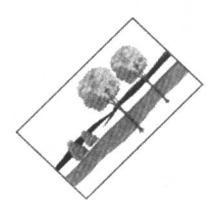
165

利用齐次坐标，求习题 3 ~ 8 中生成所给复合 2D 变换的 3×3 矩阵.

3. 平移(3,1)，然后绕原点旋转 45°.
4. 平移(−2,3)，然后将 x 坐标数乘 0. 8，y 坐标数乘 1. 2.
5. 做关于 x 轴的对称，然后绕原点旋转 30°.
6. 旋转 30°，然后做关于 x 轴的对称.
7. 绕点(6,8)旋转 60°.
8. 绕点(3,7)旋转 45°.
9. 一个 2×200 数据矩阵 $\boldsymbol{D}$ 包含了 200 个点的坐标. 使用任意两个 2×2 矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 对点进行变换时，计算所需的乘法运算量，试考虑 $\boldsymbol{A}(\boldsymbol{BD})$ 和 $(\boldsymbol{AB})\boldsymbol{D}$ 两种可能. 讨论你的结论对计算机图形计算有何意义？
10. 考虑下列 2D 几何变换：膨胀 D(它将坐标和坐标数乘同一因数)、旋转 R 以及平移 T. D 与 R 可交换吗？即 $D(R(\boldsymbol{x}))=R(D(\boldsymbol{x}))$ 是否对 $\mathbf{R}^2$ 中任意 $\boldsymbol{x}$ 都成立？D 与 T 可交换吗？R 与 T 可交换吗？
11. 计算机屏幕上图像的旋转有时用两个剪切 − 数乘变换的乘积来实现，它能加速确定图形图像按屏幕像素呈现的方式(计算机屏幕由逐行逐列的点组成，这些点被称作像素). 第一个变换垂直剪切并压缩每一列像素；第二个变换水平剪切并拉伸每一行像素. 令

$$\boldsymbol{A}_1=\begin{bmatrix}1&0&0\\\sin\varphi&\cos\varphi&0\\0&0&1\end{bmatrix},\qquad \boldsymbol{A}_2=\begin{bmatrix}\sec\varphi&-\tan\varphi&0\\0&1&0\\0&0&1\end{bmatrix}$$

证明这两个变换的合成是 $\mathbf{R}^2$ 中的一个旋转.

12. $\mathbf{R}^2$中的一个旋转通常需要4次乘法运算. 计算下面的乘积并证明：旋转矩阵可以分解成3个(仅需要1次乘法运算的)剪切变换.

$$\begin{bmatrix} 1 & -\tan\varphi/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ \sin\varphi & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -\tan\varphi/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

13. 计算机2D图形齐次坐标上的变换通常涉及下列形式的3×3矩阵：

$$\begin{bmatrix} A & \boldsymbol{p} \\ \mathbf{0}^{\mathrm{T}} & \mathbf{1} \end{bmatrix}$$

其中A是2×2矩阵，$\boldsymbol{p}\in\mathbf{R}^2$. 证明：一个这样的变换相当于在$\mathbf{R}^2$上先做一个线性变换再做一个平移. [提示：求一种适当的矩阵分解，其中仅包含分块矩阵.]

14. 证明：习题7中的变换等价于先绕原点旋转再平移$\boldsymbol{p}$. 求出$\boldsymbol{p}$.

15. $\mathbf{R}^3$中哪一个向量具有齐次坐标$\left(\frac{1}{2}, -\frac{1}{4}, \frac{1}{8}, \frac{1}{24}\right)$?

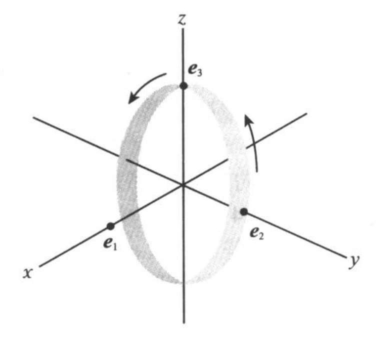
16. $(1,-2,3,4)$和$(10,-20,30,40)$是否是$\mathbf{R}^3$中同一个点的齐次坐标？为什么？

17. 给出$\mathbf{R}^3$中绕x轴旋转60°的变换所对应的4×4矩阵.(见右图.)

18. 给出$\mathbf{R}^3$中绕z轴旋转$-30°$，然后平移$\boldsymbol{p}=(5,-2,1)$的变换所对应的4×4矩阵.

19. 设S是顶点为$(4.2,1.2,4)$，$(6,4,2)$和$(2,2,6)$的三角形. 求以$(0,0,10)$为投影中心时，S在透视投影下的像.

20. 设S是顶点为$(9,3,-5)$，$(12,8,2)$和$(1.8,2.7,1)$的三角形. 求以$(0,0,10)$为投影中心时，S在透视投影下的像.



习题17的图

习题21和22讨论了计算机图形上的色彩呈现方式. 计算机屏幕上的一种颜色由三个数值(R,G,B)来编码，它们分别表示电子枪射向计算机屏幕上红、绿、蓝三色荧光点的能量大小(还可以用第四个数指定色彩的亮度或者强度).

21. [M]观察者在屏幕上实际看到的色彩要受色彩制式和屏幕上荧光点数量的影响. 因此每家计算机屏幕制造商都必须在(R,G,B)数据和国际通行的CIE色彩标准之间进行转换，CIE标准使用三原色，分别称为X，Y和Z. 针对短余辉荧光点的一类典型转换是

$$\begin{bmatrix} 0.61 & 0.29 & 0.150 \\ 0.35 & 0.59 & 0.063 \\ 0.04 & 0.12 & 0.787 \end{bmatrix}\begin{bmatrix} R \\ G \\ B \end{bmatrix}=\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

计算机程序把用CIE数据(X,Y,Z)表示的色彩信息流发送至屏幕. 求屏幕上的电子枪将这些数据转换成(R,G,B)数据的方程.

22. [M]民用电视信号发送使用向量(Y,I,Q)来描述每种颜色. 如果屏幕是黑白的，则只用到了Y(这比CIE数据能提供更好的单色图像). YIQ与"标准"RGB色彩之间的对应如下：

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix}=\begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.528 & 0.311 \end{bmatrix}\begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

(屏幕制造商需要调整矩阵元素以适应其RGB屏幕.)求将电视台发送的数据转换为电视机屏幕所要求数据的方程.

166

基础练习答案

按从右到左的顺序将三个变换对应的矩阵进行合成. 已知 $\boldsymbol{p}=(-2,6)$, $\cos(-30°)=\sqrt{3}/2$ 以及 $\sin(-30°)=-0.5$,我们有

$$\underset{\text{平移 }\boldsymbol{p}}{\begin{bmatrix}1&0&-2\\0&1&6\\0&0&1\end{bmatrix}}\underset{\text{绕原点旋转}}{\begin{bmatrix}\sqrt{3}/2&1/2&0\\-1/2&\sqrt{3}/2&0\\0&0&1\end{bmatrix}}\underset{\text{平移 }-\boldsymbol{p}}{\begin{bmatrix}1&0&2\\0&0&-6\\0&0&1\end{bmatrix}}=\begin{bmatrix}\sqrt{3}/2&1/2&\sqrt{3}-5\\-1/2&\sqrt{3}/2&-3\sqrt{3}+5\\0&0&1\end{bmatrix}$$

2.8　$\mathbf{R}^n$的子空间

本节将集中讨论由 $\mathbf{R}^n$ 中向量构成的一类重要集合——子空间. 子空间常与某个矩阵 $\boldsymbol{A}$ 相关,它们给出了方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 的有关信息. 本节中的概念和术语将在本书其余部分反复使用.[1]

167

【定义】　$\mathbf{R}^n$的一个子空间是 $\mathbf{R}^n$ 中满足下列三个性质的任意集合 H:

a. 零向量属于 H.

b. 对 H 中的任意 $\boldsymbol{u}$ 和 $\boldsymbol{v}$,和 $\boldsymbol{u}+\boldsymbol{v}$ 属于 H.

c. 对 H 中任意 $\boldsymbol{u}$ 以及任意数量 c,向量 $c\boldsymbol{u}$ 属于 H.

简言之,子空间对加法和数乘运算封闭. 在下面几个例题中将看到,第 1 章讨论的大部分向量集合都是子空间. 例如,过原点的平面是例题 1 中子空间的标准图像. 见图 2-24.

【例题 1】　若 $\boldsymbol{v}_1,\ \boldsymbol{v}_2\in\mathbf{R}^n$ 且 $H=\text{Span}\{\boldsymbol{v}_1,\ \boldsymbol{v}_2\}$,则 H 是 $\mathbf{R}^n$的子空间. 为证明这一命题,首先注意零向量属于 H(因为 $0\boldsymbol{v}_1+0\boldsymbol{v}_2$是 $\boldsymbol{v}_1$和 $\boldsymbol{v}_2$的线性组合). 现在任取 H 中两个向量,比如,

$$\boldsymbol{u}=s_1\boldsymbol{v}_1+s_2\boldsymbol{v}_2\ \text{和}\ \boldsymbol{v}=t_1\boldsymbol{v}_1+t_2\boldsymbol{v}_2$$

则

$$\boldsymbol{u}+\boldsymbol{v}=(s_1+t_1)\boldsymbol{v}_1+(s_2+t_2)\boldsymbol{v}_2$$

这就表明 $\boldsymbol{u}+\boldsymbol{v}$ 是 $\boldsymbol{v}_1$和 $\boldsymbol{v}_2$的线性组合,从而属于 H. 此外,对任意数量 c,向量 $c\boldsymbol{u}$ 属于 H,这是因为 $c\boldsymbol{u}=c(s_1\boldsymbol{v}_1+s_2\boldsymbol{v}_2)=(cs_1)\boldsymbol{v}_1+(cs_2)\boldsymbol{v}_2$. ■

若 $\boldsymbol{v}_1$非零且 $\boldsymbol{v}_2$是 $\boldsymbol{v}_1$的倍数,则 $\boldsymbol{v}_1$和 $\boldsymbol{v}_2$只能张成过原点的一条直线(见图 2-25). 所以过原点的直线是另一个子空间的例子.

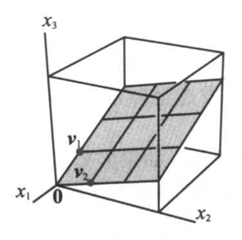


图 2-24　Span$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$是一个过原点的平面

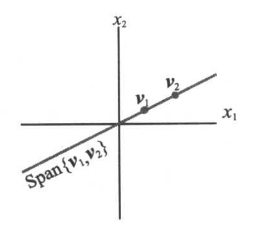


图 2-25　$\boldsymbol{v}_1\neq 0$,$\boldsymbol{v}_2=k\boldsymbol{v}_1$

1. 把 2.8 和 2.9 两节安排在这里,将允许读者暂缓学习后两章的大部分或全部内容,直接跳至第 5 章. 如果读者计划在进入第 5 章前先学习第 4 章,则可以略过这两节.

【例题2】 一条不经过原点的直线 L 不是子空间，因为按子空间的定义原点应包含在内. 此外，图 2-26 还表明 L 对加法和数乘不封闭. ■

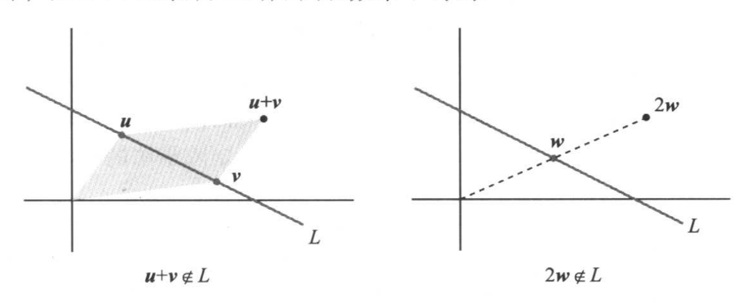


图 2-26

【例题3】 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$是 $\mathbf{R}^n$ 中向量，它们的线性组合所构成的集合是 $\mathbf{R}^n$的子空间. 这个命题的证明与例题 1 类似. 我们以后用 Span$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$表示**由 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$ 张成(或生成)的子空间**[the subspace spanned(or generated) by $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$]. ■ 168

注意 $\mathbf{R}^n$满足子空间的三条性质，因而是其自身的子空间. 另一个特殊的例子是 $\mathbf{R}^n$中仅含零向量的集合，这个集合被称作**零子空间**(zero subspace)，它同样满足子空间的条件.

2.8.1 矩阵的列空间和零空间

理论和应用中常见的 $\mathbf{R}^n$的子空间有两种类型，它们都与矩阵有关.

【定义】 矩阵 A 的**列空间**(column space)是集合 Col A，它由 A 中列的全体线性组合构成.

如果 $A=[\boldsymbol{a}_1\cdots\boldsymbol{a}_n]$，其中各列属于 $\mathbf{R}^m$，则 Col A 与 Span$\{\boldsymbol{a}_1,\ \cdots,\ \boldsymbol{a}_n\}$相等. 例题 3 表明 **$m\times n$ 矩阵的列空间是 $\mathbf{R}^m$的子空间**.

【例题4】 设 $A=\begin{bmatrix}1&-3&-4\\-4&6&-2\\-3&7&6\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}3\\3\\-4\end{bmatrix}$. 判断 $\boldsymbol{b}$ 是否属于 A 的列空间.

解： 向量 $\boldsymbol{b}$ 是 A 中列的线性组合当且仅当存在某个 $\boldsymbol{x}$ 使得 $\boldsymbol{b}$ 能写成 $A\boldsymbol{x}$，即当且仅当方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有解. 对增广矩阵$[A\quad \boldsymbol{b}]$进行行化简，

$$\begin{bmatrix}1&-3&-4&3\\-4&6&-2&3\\-3&7&6&-4\end{bmatrix}\sim\begin{bmatrix}1&-3&-4&3\\0&-6&-18&15\\0&-2&-6&5\end{bmatrix}\sim\begin{bmatrix}1&-3&-4&3\\0&-6&-18&15\\0&0&0&0\end{bmatrix}$$

我们知道 $A\boldsymbol{x}=\boldsymbol{b}$ 是相容的，于是 $\boldsymbol{b}$ 属于 Col A. ■

例题 4 的求解过程表明，当一个线性方程组写成 $A\boldsymbol{x}=\boldsymbol{b}$ 的形式时，A 的列空间是使方程组有解的全体 $\boldsymbol{b}$ 所构成的集合.

【定义】 矩阵 A 的**零空间**(null space)是集合 Nul A，它由齐次方程 $A\boldsymbol{x}=\boldsymbol{0}$ 的全体解构成.

当 $\boldsymbol{A}$ 含有 n 列时，$\boldsymbol{Ax}=\boldsymbol{0}$ 的解属于 $\mathbf{R}^n$，$\boldsymbol{A}$ 的零空间是 $\mathbf{R}^n$ 的子集．事实上，Nul $\boldsymbol{A}$ 满足 $\mathbf{R}^n$ 子空间的性质．

169

【定理12】　$m\times n$ 矩阵 $\boldsymbol{A}$ 的零空间是 $\mathbf{R}^n$ 的子空间．即含 m 个齐次方程、n 个未知量的方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解集是 $\mathbf{R}^n$ 的子空间．

证明： 零向量属于 Nul $\boldsymbol{A}$（因为 $\boldsymbol{A0}=\boldsymbol{0}$）．为了证明 Nul $\boldsymbol{A}$ 满足子空间的另外两条性质，任取 $\boldsymbol{u}$，$\boldsymbol{v}\in$ Nul $\boldsymbol{A}$，即假定 $\boldsymbol{Au}=\boldsymbol{0}$ 和 $\boldsymbol{Av}=\boldsymbol{0}$，则由矩阵乘法的性质有

$$\boldsymbol{A}(\boldsymbol{u}+\boldsymbol{v})=\boldsymbol{Au}+\boldsymbol{Av}=\boldsymbol{0}+\boldsymbol{0}=\boldsymbol{0}$$

即 $\boldsymbol{u}+\boldsymbol{v}$ 满足 $\boldsymbol{Ax}=\boldsymbol{0}$，于是 $\boldsymbol{u}+\boldsymbol{v}\in$ Nul $\boldsymbol{A}$．此外，对任意数量 c，$\boldsymbol{A}(c\boldsymbol{u})=c(\boldsymbol{Au})=c(\boldsymbol{0})=\boldsymbol{0}$，表明 $c\boldsymbol{u}\in$ Nul $\boldsymbol{A}$．■

为了判断一个给定的向量 $\boldsymbol{v}$ 是否属于 Nul $\boldsymbol{A}$，只需计算 $\boldsymbol{Av}$，看它是否为零向量．由于 Nul $\boldsymbol{A}$ 中的向量都必须满足这个约束条件，因此我们说 Nul $\boldsymbol{A}$ 被隐式定义．对比之下，Col $\boldsymbol{A}$ 列空间被显式定义，因为 Col $\boldsymbol{A}$ 中的向量可以用 $\boldsymbol{A}$ 中列的线性组合来构造．为了得到 Nul $\boldsymbol{A}$ 的一个显式定义，我们可以求解方程 $\boldsymbol{Ax}=\boldsymbol{0}$，将它的解用参数向量形式表示出来（见例题6）[1]．

2.8.2　子空间的基

由于子空间通常含有无穷多个向量，处理子空间有关问题的最好方法是研究它的有限张成集．这个集合越小，处理起来越简单．可以证明最小的张成集线性无关．

【定义】　$\mathbf{R}^n$ 中子空间 H 的**基**（basis）是 H 中能够张成 H 的线性无关集．

【例题5】　根据可逆矩阵定理，$n\times n$ 可逆矩阵的列线性无关且张成 $\mathbf{R}^n$，因而构成 $\mathbf{R}^n$ 的一组基．$n\times n$ 单位矩阵就是这样的矩阵，以 $\boldsymbol{e}_1$，…，$\boldsymbol{e}_n$ 记它的各列：

$$\boldsymbol{e}_1=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix},\ \boldsymbol{e}_2=\begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix},\ \cdots,\ \boldsymbol{e}_n=\begin{bmatrix}0\\\vdots\\0\\1\end{bmatrix}$$

170

集合 $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$ 称作 $\mathbf{R}^n$ 的**标准基**（standard basis）．见图2-27．■

下一道例题表明，将 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解集写成参数向量形式的标准方法，实际上就是求 Nul $\boldsymbol{A}$ 的一组基．整个第5章都将利用这个事实．

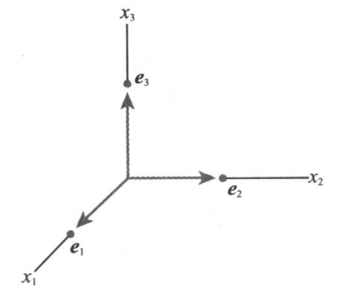


图2-27　$\mathbf{R}^3$ 的标准基

【例题6】　求下列矩阵零空间的一组基：

$$\boldsymbol{A}=\begin{bmatrix}-3&6&-1&1&-7\\1&-2&2&3&-1\\2&-4&5&8&-4\end{bmatrix}$$

解： 首先，将 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解写成参数向量形式：

---

1. Nul $\boldsymbol{A}$ 与 Col $\boldsymbol{A}$ 的异同在4.2节会做进一步讨论．

$$A \sim \begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{aligned} x_1 - 2x_2 \qquad - x_4 + 3x_5 &= 0 \\ x_3 + 2x_4 - 2x_5 &= 0 \\ 0 &= 0 \end{aligned}$$

通解是 $x_1 = 2x_2 + x_4 - 3x_5$，$x_3 = -2x_4 + 2x_5$，其中 x_2，x_4和 x_5是自由变量.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \underset{\substack{\uparrow \\ \boldsymbol{u}}}{\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}} + x_4 \underset{\substack{\uparrow \\ \boldsymbol{v}}}{\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}} + x_5 \underset{\substack{\uparrow \\ \boldsymbol{w}}}{\begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}} \tag{1}$$

$$= x_2\boldsymbol{u} + x_4\boldsymbol{v} + x_5\boldsymbol{w}$$

(1)式说明 Nul $\boldsymbol{A}$ 就是 $\boldsymbol{u}$，$\boldsymbol{v}$ 和 $\boldsymbol{w}$ 的线性组合所构成的集合，即$\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\}$生成 Nul $\boldsymbol{A}$. 事实上，由(1)可知，仅当系数 x_2，x_4和 x_5全都为零时，才有 $\mathbf{0} = x_2\boldsymbol{u} + x_4\boldsymbol{v} + x_5\boldsymbol{w}$(分别考察向量的第2，4，5个位置的元素)，因而 $\boldsymbol{u}$，$\boldsymbol{v}$ 和 $\boldsymbol{w}$ 的构造自动使得它们线性无关. 所以$\{\boldsymbol{u},\ \boldsymbol{v},\ \boldsymbol{w}\}$是 Nul $\boldsymbol{A}$ 的一组基. ■

实际上，求矩阵列空间的基比求零空间的基更容易，不过所用的方法需略加说明. 我们先来看一个简单的例子.

【例题7】 求下列矩阵列空间的一组基：

$$B = \begin{bmatrix} 1 & 0 & -3 & 5 & 0 \\ 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

解： 以 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_5$记矩阵 $\boldsymbol{B}$ 的列，注意到 $\boldsymbol{b}_3 = -3\boldsymbol{b}_1 + 2\boldsymbol{b}_2$且 $\boldsymbol{b}_4 = 5\boldsymbol{b}_1 - \boldsymbol{b}_2$. $\boldsymbol{b}_3$和 $\boldsymbol{b}_4$是主元列的线性组合，这就意味着任何 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_5$的线性组合其实都是 $\boldsymbol{b}_1$，$\boldsymbol{b}_2$和 $\boldsymbol{b}_5$的线性组合. 事实上，若 $\boldsymbol{v}$ 是 Col $\boldsymbol{B}$ 中的向量，比如说，

171

$$\boldsymbol{v} = c_1\boldsymbol{b}_1 + c_2\boldsymbol{b}_2 + c_3\boldsymbol{b}_3 + c_4\boldsymbol{b}_4 + c_5\boldsymbol{b}_5$$

则替换掉 $\boldsymbol{b}_3$和 $\boldsymbol{b}_4$后，我们可将 $\boldsymbol{v}$ 写成 $\boldsymbol{b}_1$，$\boldsymbol{b}_2$和 $\boldsymbol{b}_5$的线性组合：

$$\boldsymbol{v} = c_1\boldsymbol{b}_1 + c_2\boldsymbol{b}_2 + c_3(-3\boldsymbol{b}_1 + 2\boldsymbol{b}_2) + c_4(5\boldsymbol{b}_1 - \boldsymbol{b}_2) + c_5\boldsymbol{b}_5$$

于是$\{\boldsymbol{b}_1,\ \boldsymbol{b}_2,\ \boldsymbol{b}_5\}$张成 Col $\boldsymbol{B}$. 此外，由于 $\boldsymbol{b}_1$，$\boldsymbol{b}_2$和 $\boldsymbol{b}_5$是单位矩阵的列，因而线性无关. 所以 $\boldsymbol{B}$ 的主元列构成 Col $\boldsymbol{B}$ 的一组基. ■

例题7 中的矩阵 $\boldsymbol{B}$ 有简化阶梯形式. 为了处理一般的矩阵 $\boldsymbol{A}$，回忆 $\boldsymbol{A}$ 中列的线性相关性可以表示为：存在某个 $\boldsymbol{x}$ 使得 $\boldsymbol{Ax} = \mathbf{0}$(若某些列不出现在线性相关关系中，则它们在 $\boldsymbol{x}$ 中对应的元素是零). 当 $\boldsymbol{A}$ 行化简为阶梯形式 $\boldsymbol{B}$ 后，矩阵的列发生了变化，但方程 $\boldsymbol{Ax} = \mathbf{0}$ 和 $\boldsymbol{Bx} = \mathbf{0}$ 的解集相等. 即 $\boldsymbol{A}$ 的列和 $\boldsymbol{B}$ 的列有相同的线性相关关系.

【例题8】 可以验证矩阵

$$A=\begin{bmatrix}a_1 & a_2 & \cdots & a_5\end{bmatrix}=\begin{bmatrix}1&3&3&2&-9\\-2&-2&2&-8&2\\2&3&0&7&1\\3&4&-1&11&-8\end{bmatrix}$$

与例题7中的矩阵 $\boldsymbol{B}$ 行等价. 求 Col $\boldsymbol{A}$ 的一组基.

解: 由例题7可知, $\boldsymbol{A}$ 的主元列是第1, 2, 5列. 此外有 $\boldsymbol{b}_3=-3\boldsymbol{b}_1+2\boldsymbol{b}_2$ 和 $\boldsymbol{b}_4=5\boldsymbol{b}_1-\boldsymbol{b}_2$. 因为行变换不改变矩阵各列的线性相关关系, 我们有

$$\boldsymbol{a}_3=-3\boldsymbol{a}_1+2\boldsymbol{a}_2 \text{ 和 } \boldsymbol{a}_4=5\boldsymbol{a}_1-\boldsymbol{a}_2$$

加以验证! 根据例题7的讨论, 生成 $\boldsymbol{A}$ 的列空间时不需要用 $\boldsymbol{a}_3$ 和 $\boldsymbol{a}_4$. 此外, 由于 $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ 和 $\boldsymbol{a}_5$ 的线性相关关系必定蕴含 $\boldsymbol{b}_1$, $\boldsymbol{b}_2$ 和 $\boldsymbol{b}_5$ 的线性相关关系, 而 $\{\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_5\}$ 线性无关, 所以 $\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_5\}$ 也线性无关, 因而是 Col $\boldsymbol{A}$ 的一组基. ■

对例题8的讨论稍加改造, 可以证明下面的定理.

【定理13】 矩阵 $\boldsymbol{A}$ 的主元列构成其列空间的一组基.

警告: 注意要用 $\boldsymbol{A}$ 本身的主元列作成 Col $\boldsymbol{A}$ 的基. 阶梯形式 $\boldsymbol{B}$ 的列往往不属于 $\boldsymbol{A}$ 的列空间(例如在例题7和8中, $\boldsymbol{B}$ 中列的最后一个元素都是零, 它们不可能生成 $\boldsymbol{A}$ 的列).

172

基础练习

1. 设 $\boldsymbol{A}=\begin{bmatrix}1&-1&5\\2&0&7\\-3&-5&-3\end{bmatrix}$, $\boldsymbol{u}=\begin{bmatrix}-7\\3\\2\end{bmatrix}$. 是否有 $\boldsymbol{u}\in$ Nul $\boldsymbol{A}$? $\boldsymbol{u}\in$ Col $\boldsymbol{A}$? 说明理由.

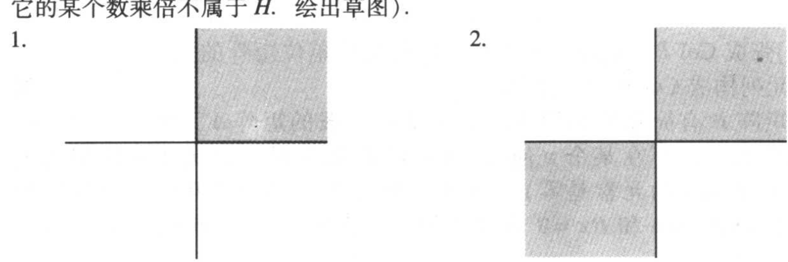
2. 已知 $\boldsymbol{A}=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}$, 分别求 Nul $\boldsymbol{A}$ 和 Col $\boldsymbol{A}$ 中的一个向量.

3. 假定 $n\times n$ 矩阵 $\boldsymbol{A}$ 可逆. 关于 Col $\boldsymbol{A}$ 你可以得到什么结论? Nul $\boldsymbol{A}$ 呢?

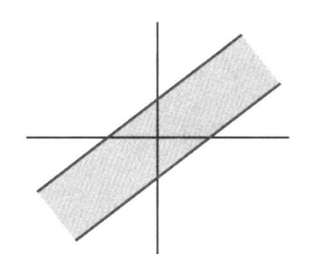
习题 2.8

习题1~4给出了 $\mathbf{R}^2$ 中的若干集合. 假设它们都包括边界线. 试就每个集合不是 $\mathbf{R}^2$ 中子空间给出明确理由(比如找出 H 中两个向量, 它们的和不属于 H, 或者找出 H 中的一个向量, 它的某个数乘倍不属于 H. 绘出草图).

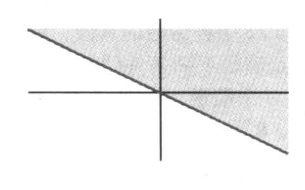
1. 2.



3.



4.



5. 设 $\boldsymbol{v}_1=\begin{bmatrix}2\\3\\-5\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}-4\\-5\\8\end{bmatrix}$，且 $\boldsymbol{w}=\begin{bmatrix}8\\2\\-9\end{bmatrix}$. 判断 $\boldsymbol{w}$ 是否属于 $\mathbf{R}^3$ 中 $\boldsymbol{v}_1$，$\boldsymbol{v}_2$生成的子空间.

6. 设 $\boldsymbol{v}_1=\begin{bmatrix}1\\-2\\4\\3\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}4\\-7\\9\\7\end{bmatrix}$，$\boldsymbol{v}_3=\begin{bmatrix}5\\-8\\6\\5\end{bmatrix}$，且 $\boldsymbol{u}=\begin{bmatrix}-4\\10\\-7\\-5\end{bmatrix}$. 判断 $\boldsymbol{u}$ 是否属于 $\mathbf{R}^4$ 中$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$生成的子空间.

7. 设 $\boldsymbol{v}_1=\begin{bmatrix}2\\-8\\6\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}-3\\8\\-7\end{bmatrix}$，$\boldsymbol{v}_3=\begin{bmatrix}-4\\6\\-7\end{bmatrix}$，$\boldsymbol{p}=\begin{bmatrix}6\\-10\\11\end{bmatrix}$，且 $A=[\boldsymbol{v}_1\quad \boldsymbol{v}_2\quad \boldsymbol{v}_3]$.

a. $\{\boldsymbol{v}_1，\boldsymbol{v}_2，\boldsymbol{v}_3\}$中有多少个向量？

b. Col $\boldsymbol{A}$ 中有多少个向量？

c. $\boldsymbol{p}$ 是否属于 Col $\boldsymbol{A}$？为什么？

8. 设 $\boldsymbol{v}_1=\begin{bmatrix}-3\\0\\6\end{bmatrix}$，$\boldsymbol{v}_2=\begin{bmatrix}-2\\2\\3\end{bmatrix}$，$\boldsymbol{v}_3=\begin{bmatrix}0\\-6\\3\end{bmatrix}$，且 $\boldsymbol{p}=\begin{bmatrix}1\\14\\-9\end{bmatrix}$. 判断 $\boldsymbol{p}$ 是否属于 Col $\boldsymbol{A}$，其中 $A=[\boldsymbol{v}_1\quad \boldsymbol{v}_2\quad \boldsymbol{v}_3]$.

173

9. 设 $\boldsymbol{A}$ 和 $\boldsymbol{p}$ 同习题 7，判断 $\boldsymbol{p}$ 是否属于 Nul $\boldsymbol{A}$.

10. 设 $\boldsymbol{u}=(-2,3,1)$，$\boldsymbol{A}$ 同习题 8，判断 $\boldsymbol{u}$ 是否属于 Nul $\boldsymbol{A}$.

在习题 11 和 12 中，求整数 p 和 q，使得 Nul $\boldsymbol{A}$ 是 $\mathbf{R}^p$的子空间，Col $\boldsymbol{A}$ 是 $\mathbf{R}^q$的子空间.

11. $A=\begin{bmatrix}3&2&1&-5\\-9&-4&1&7\\9&2&-5&1\end{bmatrix}$

12. $A=\begin{bmatrix}1&2&3\\4&5&7\\-5&-1&0\\2&7&11\end{bmatrix}$

13. 对于习题 11 中的 $\boldsymbol{A}$，分别求 Nul $\boldsymbol{A}$ 和 Col $\boldsymbol{A}$ 中的一个非零向量.

14. 对于习题 12 中的 $\boldsymbol{A}$，分别求 Nul $\boldsymbol{A}$ 和 Col $\boldsymbol{A}$ 中的一个非零向量.

判断习题 15 ~ 20 中哪些集合构成 $\mathbf{R}^2$或 $\mathbf{R}^3$的基，说明理由.

15. $\begin{bmatrix}5\\-2\end{bmatrix}$，$\begin{bmatrix}10\\-3\end{bmatrix}$

16. $\begin{bmatrix}-4\\6\end{bmatrix}$，$\begin{bmatrix}2\\-3\end{bmatrix}$

17. $\begin{bmatrix}0\\1\\-2\end{bmatrix}$，$\begin{bmatrix}5\\-7\\4\end{bmatrix}$，$\begin{bmatrix}6\\3\\5\end{bmatrix}$

18. $\begin{bmatrix}1\\1\\-2\end{bmatrix}$，$\begin{bmatrix}-5\\-1\\2\end{bmatrix}$，$\begin{bmatrix}7\\0\\-5\end{bmatrix}$

19. $\begin{bmatrix}3\\-8\\1\end{bmatrix}$，$\begin{bmatrix}6\\2\\-5\end{bmatrix}$

20. $\begin{bmatrix}1\\-6\\-7\end{bmatrix}$，$\begin{bmatrix}3\\-4\\7\end{bmatrix}$，$\begin{bmatrix}-2\\7\\5\end{bmatrix}$，$\begin{bmatrix}0\\8\\9\end{bmatrix}$

判断习题21和22中各命题的真假，说明理由.

21. a. $\mathbf{R}^n$的子空间是满足下列条件的任意集合 H：(i)零向量属于 H，(ii)$\boldsymbol{u}$，$\boldsymbol{v}$ 和 $\boldsymbol{u}+\boldsymbol{v}$ 都属于 H，(iii)c 是一个数量，$c\boldsymbol{u}$ 属于 H.
    b. 若 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p \in \mathbf{R}^n$，则 Span$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$与矩阵$[\boldsymbol{v}_1\cdots\boldsymbol{v}_p]$的列空间相同.
    c. 含 m 个齐次方程、n 个未知量的方程组的解集是 $\mathbf{R}^m$的子空间.
    d. $n\times n$ 可逆矩阵的列构成 $\mathbf{R}^n$的一组基.
    e. 行变换不改变矩阵各列的线性相关关系.
22. a. $\mathbf{R}^n$的子集 H 是子空间，如果零向量属于 H.
    b. 给定 $\mathbf{R}^n$中向量 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$，它们的全体线性组合所构成的集合是 $\mathbf{R}^n$的子空间.
    c. $m\times n$ 矩阵的零空间是 $\mathbf{R}^n$的子空间.
    d. 矩阵 $\boldsymbol{A}$ 的列空间是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解集.
    e. 若 $\boldsymbol{B}$ 是矩阵 $\boldsymbol{A}$ 的阶梯形式，则 $\boldsymbol{B}$ 的主元列构成 Col $\boldsymbol{A}$ 的一组基.

习题23～26中给出了矩阵 $\boldsymbol{A}$ 及其阶梯形式. 分别求 Col $\boldsymbol{A}$ 和 Nul $\boldsymbol{A}$ 的一组基.

23. $\boldsymbol{A}=\begin{bmatrix}4&5&9&-2\\6&5&1&12\\3&4&8&-3\end{bmatrix}\sim\begin{bmatrix}1&2&6&-5\\0&1&5&-6\\0&0&0&0\end{bmatrix}$

24. $\boldsymbol{A}=\begin{bmatrix}-3&9&-2&-7\\2&-6&4&8\\3&-9&-2&2\end{bmatrix}\sim\begin{bmatrix}1&-3&6&9\\0&0&4&5\\0&0&0&0\end{bmatrix}$

25. $\boldsymbol{A}=\begin{bmatrix}1&4&8&-3&-7\\-1&2&7&3&4\\-2&2&9&5&5\\3&6&9&-5&-2\end{bmatrix}\sim\begin{bmatrix}1&4&8&0&5\\0&2&5&0&-1\\0&0&0&1&4\\0&0&0&0&0\end{bmatrix}$

26. $\boldsymbol{A}=\begin{bmatrix}3&-1&7&3&9\\-2&2&-2&7&5\\-5&9&3&3&4\\-2&6&6&3&7\end{bmatrix}\sim\begin{bmatrix}3&-1&7&0&6\\0&2&4&0&3\\0&0&0&1&1\\0&0&0&0&0\end{bmatrix}$

27. 构造 3×3 矩阵 $\boldsymbol{A}$ 及非零向量 $\boldsymbol{b}$，使 $\boldsymbol{b}$ 属于 Col $\boldsymbol{A}$，但 $\boldsymbol{b}$ 不等于 $\boldsymbol{A}$ 的任何一列.
28. 构造 3×3 矩阵 $\boldsymbol{A}$ 及向量 $\boldsymbol{b}$，使 $\boldsymbol{b}$ 不属于 Col $\boldsymbol{A}$.
29. 构造 3×3 矩阵 $\boldsymbol{A}$ 及非零向量 $\boldsymbol{b}$，使 $\boldsymbol{b}$ 属于 Nul $\boldsymbol{A}$.
30. 假定矩阵 $\boldsymbol{A}=[\boldsymbol{a}_1\cdots\boldsymbol{a}_p]$各列线性无关. 解释$\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_p\}$为什么是 Col $\boldsymbol{A}$ 的一组基.

174

尽可能全面地回答习题31～36中的问题，并说明理由.

31. 假设 $\boldsymbol{F}$ 是一个 5×5 矩阵，其列空间不等于 $\mathbf{R}^5$，那么关于 Nul $\boldsymbol{F}$ 你能得出什么结论?
32. 如果 $\boldsymbol{R}$ 是一个 6×6 矩阵，Nul $\boldsymbol{R}$ 不是零子空间，那么关于 Col $\boldsymbol{R}$ 你能得出什么结论?
33. 如果 $\boldsymbol{Q}$ 是一个 4×4 矩阵，且 Col $\boldsymbol{Q}=\mathbf{R}^4$，那么关于方程 $\boldsymbol{Qx}=\boldsymbol{b}$，$\boldsymbol{b}\in\mathbf{R}^4$的解，你能得出什么结论?
34. 如果 $\boldsymbol{P}$ 是一个 5×5 矩阵，Nul $\boldsymbol{P}$ 是零子空间，那么关于方程 $\boldsymbol{Px}=\boldsymbol{b}$，$\boldsymbol{b}\in\mathbf{R}^5$的解，你能得出什么结论?
35. 当 $\boldsymbol{B}$ 是各列线性无关的 5×4 矩阵时，关于 Nul $\boldsymbol{B}$ 你能得出什么结论?
36. 当 $m\times n$ 矩阵 $\boldsymbol{A}$ 的列构成 $\mathbf{R}^m$的一组基时，关于 $\boldsymbol{A}$ 的形状你能得出什么结论?

构造习题37和38中矩阵 $\boldsymbol{A}$ 的列空间和零空间的基，并加以验证.

37. $A=\begin{bmatrix} 3 & -5 & 0 & -1 & 3 \\ -7 & 9 & -4 & 9 & -11 \\ -5 & 7 & -2 & 5 & -7 \\ 3 & -7 & -3 & 4 & 0 \end{bmatrix}$

38. $A=\begin{bmatrix} 5 & 2 & 0 & -8 & -8 \\ 4 & 1 & 2 & -8 & -9 \\ 5 & 1 & 3 & 5 & 19 \\ -8 & -5 & 6 & 8 & 5 \end{bmatrix}$

基础练习答案

1. 为判断 $\boldsymbol{u}$ 是否属于 Nul A，只需计算

$$A\boldsymbol{u}=\begin{bmatrix} 1 & -1 & 5 \\ 2 & 0 & 7 \\ -3 & -5 & -3 \end{bmatrix}\begin{bmatrix} -7 \\ 3 \\ 2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

结果表明 $\boldsymbol{u}\in$ Nul A. 确定 $\boldsymbol{u}$ 是否属于 Col A 还需进一步讨论. 将增广矩阵$[A\quad \boldsymbol{u}]$化简成阶梯形式，以判断方程 $A\boldsymbol{x}=\boldsymbol{u}$ 是否相容：

$$\begin{bmatrix} 1 & -1 & 5 & -7 \\ 2 & 0 & 7 & 3 \\ -3 & -5 & -3 & 2 \end{bmatrix}\sim\begin{bmatrix} 1 & -1 & 5 & -7 \\ 0 & 2 & -3 & 17 \\ 0 & -8 & 12 & -19 \end{bmatrix}\sim\begin{bmatrix} 1 & -1 & 5 & -7 \\ 0 & 2 & -3 & 17 \\ 0 & 0 & 0 & 49 \end{bmatrix}$$

方程 $A\boldsymbol{x}=\boldsymbol{u}$ 无解，所以 $\boldsymbol{u}$ 不属于 Col A.

2. 与基础练习 1 形成对照，求 Nul A 的一个向量比判断某个给定的向量是否属于 Nul A 要困难. 不过，由于 A 已经是简化阶梯形式，方程 $A\boldsymbol{x}=\boldsymbol{0}$ 表明：如果 $\boldsymbol{x}=(x_1,x_2,x_3)$，则 $x_2=0$，$x_3=0$ 且 x_1 是自由变量. 于是 $\boldsymbol{v}=(1,0,0)$ 是 Nul A 的一组基. 求 Col A 的一个向量很容易，因为 A 的每一列都属于 Col A. 特别地，向量 $\boldsymbol{v}$ 同时属于 Nul A 和 Col A. 对于大多数 $n\times n$ 矩阵来说，$\mathbf{R}^n$ 中的零向量是唯一一个同时属于 Nul A 和 Col A 的向量.

3. 若 A 可逆，则根据可逆矩阵定理，A 的列张成 $\mathbf{R}^n$. 由定义知，任意矩阵的列都张成其列空间，所以 Col A 就是 $\mathbf{R}^n$ 的全体. 用符号表示为 Col $A=\mathbf{R}^n$. 此外，由于 A 可逆，方程 $A\boldsymbol{x}=\boldsymbol{0}$ 有唯一平凡解. 这就意味着 Nul A 是零子空间. 用符号表示为 Nul $A=\{\boldsymbol{0}\}$.

175

2.9 维数和秩

本节首先介绍坐标系的概念，仍然继续前文关于子空间及其基的讨论. 下面的定义和例题将用到一个新概念——维数，至少对 $\mathbf{R}^3$ 的子空间来说，这是一个极为自然的概念.

2.9.1 坐标系

我们通常选取子空间 H 的一组基，而不仅仅是一个张成集，其主要原因是 H 中的向量能唯一表示成基向量的线性组合. 为证明这一点，设 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p\}$是 H 的一组基，H 中向量 $\boldsymbol{x}$ 可由下列两种方式生成：

$$\boldsymbol{x}=c_1\boldsymbol{b}_1+\cdots+c_p\boldsymbol{b}_p \text{ 和 } \boldsymbol{x}=d_1\boldsymbol{b}_1+\cdots+d_p\boldsymbol{b}_p \tag{1}$$

两式相减可得

$$\boldsymbol{0}=\boldsymbol{x}-\boldsymbol{x}=(c_1-d_1)\boldsymbol{b}_1+\cdots+(c_p-d_p)\boldsymbol{b}_p \tag{2}$$

由于 $\mathcal{B}$ 线性无关，(2)中的系数必定为零. 即对 $1\leqslant j\leqslant p$ 有 $c_j=d_j$，这就表明(1)中的两种表示实际上是相同的.

【定义】 设集合 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p\}$是子空间 H 的一组基. 对于任意 $\boldsymbol{x}\in H$，**$\boldsymbol{x}$ 关于基 $\mathcal{B}$ 的坐标**(coordinates of $\boldsymbol{x}$ relative to the basis $\mathcal{B}$)是满足 $\boldsymbol{x}=c_1\boldsymbol{b}_1+\cdots+c_p\boldsymbol{b}_p$ 的系数 $c_1,\cdots,c_p$，$\mathbf{R}^p$ 中的向量

$$[\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$$

称为 $\boldsymbol{x}$(关于基 $\mathcal{B}$ 的)坐标向量[coordinate vector of $\boldsymbol{x}$(relative to $\mathcal{B}$)]**或 $\boldsymbol{x}$ 的 $\mathcal{B}$-坐标向量**($\mathcal{B}$-coordinate vector of $\boldsymbol{x}$). [1]

【例题1】　设 $\boldsymbol{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $\boldsymbol{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$, 且 $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$. 则由 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 线性无关可知，$\mathcal{B}$ 是 $H = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ 的基. 判断 $\boldsymbol{x}$ 是否属于 H，若属于，求 $\boldsymbol{x}$ 关于 $\mathcal{B}$ 的坐标向量.

解： 若 $\boldsymbol{x} \in H$，则下列向量方程是相容的：

$$c_1 \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$$

176 数量 c_1 和 c_2 若存在，就是 $\boldsymbol{x}$ 的 $\mathcal{B}$-坐标. 经过行变换有

$$\begin{bmatrix} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

于是 $c_1 = 2$, $c_2 = 3$, 且 $[\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. 基 $\mathcal{B}$ 确定了 H 上的一个“坐标系”，它可以用图2-28中的网格来表示. ■

注意，尽管 H 中的点属于 $\mathbf{R}^3$，但是它们由 $\mathbf{R}^2$ 中的 $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$-坐标向量就已完全确定. 图2-28中平面上的网格使得 H 看起来和 $\mathbf{R}^2$ 一样. 对应 $\boldsymbol{x} \mapsto [\boldsymbol{x}]_{\mathcal{B}}$ 是 H 和 $\mathbf{R}^2$ 之间保持线性组合的一对一对应. 我们将这样的对应称为同构，并且说 H 与 $\mathbf{R}^2$ 同构.

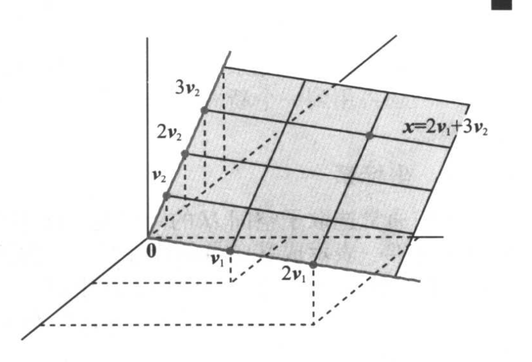


图2-28　$\mathbf{R}^3$ 中平面 H 上的一个坐标系

一般地，如果 $\mathcal{B} = \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_p\}$ 是 H 的一组基，则一对一映射 $\boldsymbol{x} \mapsto [\boldsymbol{x}]_{\mathcal{B}}$ 使得 H 无论就外观还是内部结构都与 $\mathbf{R}^p$ 一样(尽管 H 中的向量本身可能包含多于 p 个元素). 4.4节将做详细讨论.

2.9.2　子空间的维数

可以证明，如果子空间 H 有一组基包含 p 个向量，则 H 的每组基都恰好包含 p

---

1. 把 $\mathcal{B}$ 的元素列举出来十分必要，因为 $[\boldsymbol{x}]_{\mathcal{B}}$ 中的元素的次序依赖于 $\mathcal{B}$ 中向量的次序.

个向量(见习题27和28). 于是下列定义是合理的.

【定义】 非零子空间 H 的**维数**(dimension), 记作 $\dim H$, 是 H 的任意一组基所含向量的数目. 零子空间$\{\mathbf{0}\}$的维数定义成零.[1]

177

$\mathbb{R}^n$空间的维数是 n. $\mathbb{R}^n$的每组基都包含 n 个向量. $\mathbb{R}^3$中经过$\mathbf{0}$的平面是2维的, 经过$\mathbf{0}$的直线是1维的.

【例题2】 回忆2.8节例题6中矩阵 A 的零空间有一组含3个向量的基. 所以 $\mathrm{Nul}\,A$ 的维数是3. 注意, 该例中每个基向量都对应于方程 $A\mathbf{x}=\mathbf{0}$ 的一个自由变量. 我们常常用这种方式来构造基. 所以, 为了求 $\mathrm{Nul}\,A$ 的维数, 只需数出 $A\mathbf{x}=\mathbf{0}$ 中自由变量的个数. ■

【定义】 矩阵 A 的**秩**(rank), 记作 $\mathrm{rank}\,A$, 是 A 的列空间的维数.

因为 A 的主元列构成 $\mathrm{Col}\,A$ 的一组基, A 的秩恰好是 A 中主元列的数目.

【例题3】 确定下列矩阵的秩:

$$A=\begin{bmatrix}2 & 5 & -3 & -4 & 8\\ 4 & 7 & -4 & -3 & 9\\ 6 & 9 & -5 & 2 & 4\\ 0 & -9 & 6 & 5 & -6\end{bmatrix}$$

解: 将 A 化简为阶梯形式:

$$A\sim\begin{bmatrix}2 & 5 & -3 & -4 & 8\\ 0 & -3 & 2 & 5 & -7\\ 0 & -6 & 4 & 14 & -20\\ 0 & -9 & 6 & 5 & -6\end{bmatrix}\sim\cdots\sim\begin{bmatrix}2 & 5 & -3 & -4 & 8\\ 0 & -3 & 2 & 5 & -7\\ 0 & 0 & 0 & 4 & -6\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

主元列 (第1、2、4列)

矩阵 A 有3个主元列, 所以 $\mathrm{rank}\,A=3$. ■

例题3中的行化简表明 $A\mathbf{x}=\mathbf{0}$ 有两个自由变量, 因为 A 的五列中有两个非主元列(非主元列对应于 $A\mathbf{x}=\mathbf{0}$ 中的自由变量). 由于主元列的数目与非主元列的数目之和恰好是总列数, 所以 $\mathrm{Col}\,A$ 与 $\mathrm{Nul}\,A$ 的维数之间存在下面有用的关系(详见4.6节的秩定理).

【定理14】 秩定理

若矩阵 A 含 n 列, 则 $\mathrm{rank}\,A+\dim\mathrm{Nul}\,A=n$.

下面的定理在应用中非常重要, 第5、6章中也需要用到. 若考虑与$\mathbb{R}^p$同构的 p 维子空间, 这个定理看起来十分合理(4.5节给出了它的证明). 可逆矩阵定理表明, $\mathbb{R}^p$中的 p 个向量线性无关当且仅当它们能够张成$\mathbb{R}^p$.

178

【定理15】 基定理

设 H 是$\mathbb{R}^p$的一个 p 维子空间. H 中任何恰含 p 个元素的线性无关集都是 H 的基. H 中任何包含 p 个元素且能张成 H 的集合都是 H 的基.

2.9.3 秩与可逆矩阵定理

有了与矩阵相关的各种向量空间的概念, 我们可以为可逆矩阵定理添加几个新的命题. 沿用2.3节中原始定理的编号, 这些命题列举如下.

---

1. 零子空间没有基(因为零向量本身构成一个线性相关集).

【定理16】 可逆矩阵定理(续)

设 $\boldsymbol{A}$ 是一个 $n\times n$ 矩阵. 则下列每一个命题都等价于 $\boldsymbol{A}$ 是可逆矩阵.

(m) $\boldsymbol{A}$ 的列构成 $\mathbf{R}^n$ 的一组基.

(n) $\mathrm{Col}\,\boldsymbol{A}=\mathbf{R}^n$

(o) $\dim\mathrm{Col}\,\boldsymbol{A}=n$

(p) $\mathrm{rank}\,\boldsymbol{A}=n$

(q) $\mathrm{Nul}\,\boldsymbol{A}=\{\boldsymbol{0}\}$

(r) $\dim\mathrm{Nul}\,\boldsymbol{A}=0$

证明: 命题(m)等价于(e)和(h), 后两者分别关于线性无关和张成. 其余五个命题与2.3节中定理8的命题有下面的蕴含关系链:

$$(g)\Rightarrow(n)\Rightarrow(o)\Rightarrow(p)\Rightarrow(r)\Rightarrow(q)\Rightarrow(d)$$

命题(g)说的是方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 对任意 $\boldsymbol{b}\in\mathbf{R}^n$ 都至少有一个解, 它蕴含(n), 因为 $\mathrm{Col}\,\boldsymbol{A}$ 恰好是使方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 相容的全体 $\boldsymbol{b}$ 所构成的集合. (n)⇒(o)⇒(p)的蕴含关系可以由维数和秩的定义得到. 如果 $\boldsymbol{A}$ 的秩等于列数, 那么根据秩定理有 $\dim\mathrm{Nul}\,\boldsymbol{A}=0$, 所以 $\mathrm{Nul}\,\boldsymbol{A}=\{\boldsymbol{0}\}$. 于是(p)⇒(r)⇒(q). 此外, 由(q)可推知方程 $\boldsymbol{Ax}=\boldsymbol{0}$ 仅有平凡解, 即命题(d). 因为已知(d)和(g)都等价于 $\boldsymbol{A}$ 可逆这一命题, 定理证毕. ■

179

注记

本书讨论的许多算法有助于读者理解概念, 也适合进行简单的手算. 不过, 对于实际生活中大规模的问题, 这些算法往往不可取.

秩的判定就是一个很好的例子. 将矩阵化简成阶梯形式, 然后数出主元列, 看起来非常容易, 但除非在元素明确指定的矩阵上施行精确的算法, 否则行运算可能会改变矩阵的显秩. 例如, 如果矩阵 $\begin{bmatrix}5&7\\5&x\end{bmatrix}$ 中的 x 没有被当作7精确地存储在计算机中, 则该矩阵的秩可能是1也可能是2, 这取决于计算机是否把 $x-7$ 当成零.

在实际应用中, 矩阵 $\boldsymbol{A}$ 的有效秩经常由 $\boldsymbol{A}$ 的奇异值分解来确定, 我们将在7.4节中加以讨论.

基础练习

1. 确定 $\mathbf{R}^3$ 中由 $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ 和 $\boldsymbol{v}_3$ 张成的子空间 H 的维数. (首先求 H 的一组基.)

$$\boldsymbol{v}_1=\begin{bmatrix}2\\-8\\6\end{bmatrix},\quad \boldsymbol{v}_2=\begin{bmatrix}3\\-7\\-1\end{bmatrix},\quad \boldsymbol{v}_3=\begin{bmatrix}-1\\6\\-7\end{bmatrix}$$

2. 考虑 $\mathbf{R}^2$ 的基 $\mathcal{B}=\left\{\begin{bmatrix}1\\0.2\end{bmatrix},\begin{bmatrix}0.2\\1\end{bmatrix}\right\}$. 若 $[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}3\\2\end{bmatrix}$, 则 $\boldsymbol{x}$ 是多少?

3. $\mathbf{R}^3$ 是否可能包含一个4维子空间? 试加以解释.

习题 2.9

在习题1和2中, 求由所给向量 $[\boldsymbol{x}]_{\mathcal{B}}$ 以及基 $\mathcal{B}$ 确定的向量. 仿照基础练习2的解答, 将

答案绘成草图.

1. $\mathcal{B}=\left\{\begin{bmatrix}1\\1\end{bmatrix},\ \begin{bmatrix}2\\-1\end{bmatrix}\right\}$, $[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}3\\2\end{bmatrix}$　　2. $\mathcal{B}=\left\{\begin{bmatrix}-2\\1\end{bmatrix},\ \begin{bmatrix}3\\1\end{bmatrix}\right\}$, $[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}-1\\3\end{bmatrix}$

习题 3 ~ 6 中向量 $\boldsymbol{x}$ 属于子空间 H，H 有一组基为 $\mathcal{B}=\{\boldsymbol{b}_1,\ \boldsymbol{b}_2\}$. 求 $\boldsymbol{x}$ 的 $\mathcal{B}$-坐标向量.

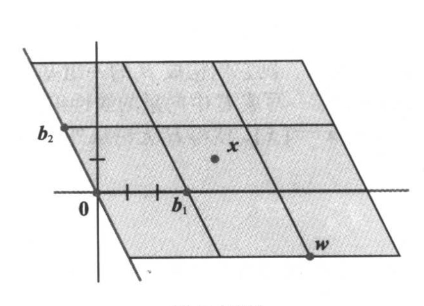
3. $\boldsymbol{b}_1=\begin{bmatrix}1\\-4\end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix}-2\\7\end{bmatrix}$, $\boldsymbol{x}=\begin{bmatrix}-3\\7\end{bmatrix}$　　4. $\boldsymbol{b}_1=\begin{bmatrix}1\\-3\end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix}-3\\5\end{bmatrix}$, $\boldsymbol{x}=\begin{bmatrix}-7\\5\end{bmatrix}$

5. $\boldsymbol{b}_1=\begin{bmatrix}1\\5\\-3\end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix}-3\\-7\\5\end{bmatrix}$, $\boldsymbol{x}=\begin{bmatrix}4\\10\\-7\end{bmatrix}$　　6. $\boldsymbol{b}_1=\begin{bmatrix}-3\\1\\-4\end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix}7\\5\\-6\end{bmatrix}$, $\boldsymbol{x}=\begin{bmatrix}11\\0\\7\end{bmatrix}$

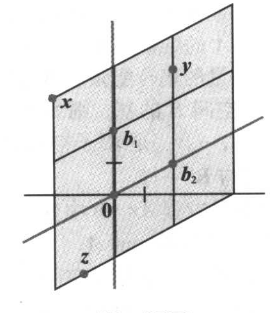
7. 设 $\boldsymbol{b}_1=\begin{bmatrix}3\\0\end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix}-1\\2\end{bmatrix}$, $\boldsymbol{w}=\begin{bmatrix}7\\-2\end{bmatrix}$, $\boldsymbol{x}=\begin{bmatrix}4\\1\end{bmatrix}$，且 $\mathcal{B}=\{\boldsymbol{b}_1,\ \boldsymbol{b}_2\}$. 画图估计 $[\boldsymbol{w}]_{\mathcal{B}}$ 和 $[\boldsymbol{x}]_{\mathcal{B}}$. 用估计的 $[\boldsymbol{x}]_{\mathcal{B}}$ 和 $\{\boldsymbol{b}_1,\ \boldsymbol{b}_2\}$ 计算 $\boldsymbol{x}$，以此验证你对 $[\boldsymbol{x}]_{\mathcal{B}}$ 的估计.

180

8. 设 $\boldsymbol{b}_1=\begin{bmatrix}0\\2\end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix}2\\1\end{bmatrix}$, $\boldsymbol{x}=\begin{bmatrix}-2\\3\end{bmatrix}$, $\boldsymbol{y}=\begin{bmatrix}2\\4\end{bmatrix}$, $\boldsymbol{z}=\begin{bmatrix}-1\\-2.5\end{bmatrix}$，且 $\mathcal{B}=\{\boldsymbol{b}_1,\ \boldsymbol{b}_2\}$. 画图估计 $[\boldsymbol{x}]_{\mathcal{B}}$，$[\boldsymbol{y}]_{\mathcal{B}}$ 和 $[\boldsymbol{z}]_{\mathcal{B}}$. 用估计的 $[\boldsymbol{y}]_{\mathcal{B}}$，$[\boldsymbol{z}]_{\mathcal{B}}$ 和 $\{\boldsymbol{b}_1,\ \boldsymbol{b}_2\}$ 计算 $\boldsymbol{y}$，$\boldsymbol{z}$，以此验证你对 $[\boldsymbol{y}]_{\mathcal{B}}$ 和 $[\boldsymbol{z}]_{\mathcal{B}}$ 的估计.



第 7 题图



第 8 题图

习题 9 ~ 12 中给出了矩阵 $\boldsymbol{A}$ 及其阶梯形式. 分别求 Col $\boldsymbol{A}$ 和 Nul $\boldsymbol{A}$ 的一组基，然后确定这些子空间的维数.

9. $\boldsymbol{A}=\begin{bmatrix}1&-3&2&-4\\-3&9&-1&5\\2&-6&4&-3\\-4&12&2&7\end{bmatrix}\sim\begin{bmatrix}1&-3&2&-4\\0&0&5&-7\\0&0&0&5\\0&0&0&0\end{bmatrix}$

10. $\boldsymbol{A}=\begin{bmatrix}1&-2&9&5&4\\1&-1&6&5&-3\\-2&0&-6&1&-2\\4&1&9&1&-9\end{bmatrix}\sim\begin{bmatrix}1&-2&9&5&4\\0&1&-3&0&-7\\0&0&0&1&-2\\0&0&0&0&0\end{bmatrix}$

11. $A=\begin{bmatrix}1&2&-5&0&-1\\2&5&-8&4&3\\-3&-9&9&-7&-2\\3&10&-7&11&7\end{bmatrix}\sim\begin{bmatrix}1&2&-5&0&-1\\0&1&2&4&5\\0&0&0&1&2\\0&0&0&0&0\end{bmatrix}$

12. $A=\begin{bmatrix}1&2&-4&3&3\\5&10&-9&-7&8\\4&8&-9&-2&7\\-2&-4&5&0&-6\end{bmatrix}\sim\begin{bmatrix}1&2&-4&3&3\\0&0&1&-2&0\\0&0&0&0&-5\\0&0&0&0&0\end{bmatrix}$

在习题 13 ~ 14 中，求给定向量所张成子空间的一组基. 子空间的维数是多少？

13. $\begin{bmatrix}1\\-3\\2\\-4\end{bmatrix},\begin{bmatrix}-3\\9\\-6\\12\end{bmatrix},\begin{bmatrix}2\\-1\\4\\2\end{bmatrix},\begin{bmatrix}-4\\5\\-3\\7\end{bmatrix}$　14. $\begin{bmatrix}1\\-1\\-2\\5\end{bmatrix},\begin{bmatrix}2\\-3\\-1\\6\end{bmatrix},\begin{bmatrix}0\\2\\-6\\8\end{bmatrix},\begin{bmatrix}-1\\4\\-7\\7\end{bmatrix},\begin{bmatrix}3\\-8\\9\\-5\end{bmatrix}$

15. 设 3×5 矩阵 $\boldsymbol{A}$ 有 3 个主元列，是否有 Col $\boldsymbol{A}=\mathbf{R}^3$？Nul $\boldsymbol{A}=\mathbf{R}^2$ 呢？说明理由.
16. 设 4×7 矩阵 $\boldsymbol{A}$ 有 3 个主元列，是否有 Col $\boldsymbol{A}=\mathbf{R}^3$？Nul $\boldsymbol{A}$ 的维数是多少？说明理由.

判断习题 17 和 18 中各命题的真假，并说明理由. 这里 $\boldsymbol{A}$ 是一个 $m\times n$ 矩阵.

17. a. 如果 $\mathcal{B}=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是子空间 H 的基，且 $\boldsymbol{x}=c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p$，则 $c_1,\cdots,c_p$ 是 $\boldsymbol{x}$ 关于基 $\mathcal{B}$ 的坐标.
    b. $\mathbf{R}^n$ 中的每条直线都是 $\mathbf{R}^n$ 的一个 1 维子空间.
    c. Col $\boldsymbol{A}$ 的维数是 $\boldsymbol{A}$ 中主元列的数目.
    d. Col $\boldsymbol{A}$ 和 Nul $\boldsymbol{A}$ 的维数之和等于 $\boldsymbol{A}$ 的列数.
    e 如果含 p 个向量的集合能张成 $\mathbf{R}^n$ 的一个 p 维子空间 H，则它们构成 H 的一组基.

181

18. a. 如果 $\mathcal{B}$ 是子空间 H 的基，则 H 中的任意向量都能唯一写成 $\mathcal{B}$ 中向量的线性组合.
    b. 如果 $\mathcal{B}=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是 $\mathbf{R}^n$ 子空间 H 的基，则对应 $\boldsymbol{x}\mapsto[\boldsymbol{x}]_{\mathcal{B}}$ 使得 H 无论从外观还是内部结构上看都与 $\mathbf{R}^p$ 一样.
    c. Nul $\boldsymbol{A}$ 的维数是方程 $\boldsymbol{Ax}=\boldsymbol{0}$ 中变量的个数.
    d. $\boldsymbol{A}$ 的列空间的维数是 rank $\boldsymbol{A}$.
    e. 如果 H 是 $\mathbf{R}^n$ 的一个 p 维子空间，则 H 中含有 p 个向量的线性无关集是 H 的基.

在习题 19 ~ 24 中，证明你的答案或者构造.

19. 如果 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解构成的子空间有一组含 3 个向量的基，且 $\boldsymbol{A}$ 是 5×7 矩阵，那么 $\boldsymbol{A}$ 的秩是多少？
20. 若一个 4×5 矩阵的零空间是 3 维的，则该矩阵的秩是多少？
21. 若一个 7×6 矩阵的秩是 4，则 $\boldsymbol{Ax}=\boldsymbol{0}$ 解空间的维数是多少？
22. 证明：$\mathbf{R}^n$ 的集合 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_5\}$ 线性相关，如果 dim Span$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_5\}=4$.
23. 如果可能，构造一个 3×4 矩阵 $\boldsymbol{A}$，使得 dim Nul $\boldsymbol{A}=2$ 且 dim Col $\boldsymbol{A}=2$.
24. 构造一个秩为 1 的 4×3 矩阵.
25. 设 $\boldsymbol{A}$ 是一个 $n\times p$ 矩阵，其列空间维数为 p. 解释为什么 $\boldsymbol{A}$ 的列一定线性无关？
26. 假设矩阵 $\boldsymbol{A}$ 的第 1、3、5、6 列线性无关(不必要求是主元列)，且 $\boldsymbol{A}$ 的秩是 4. 解释为什么这 4 列一定构成 $\boldsymbol{A}$ 的列空间的一组基？
27. 假设向量 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p$ 张成子空间 W，令 $\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_q\}$ 是 W 中包含多于 p 个向量的任何集合. 补充下列步骤的细节，以证明 $\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_q\}$ 是线性相关的. 首先，令 $\boldsymbol{B}=[\boldsymbol{b}_1\ \cdots\ \boldsymbol{b}_p]$ 和 $\boldsymbol{A}=[\boldsymbol{a}_1\ \cdots\ \boldsymbol{a}_q]$.
    a. 解释为什么对任意向量 $\boldsymbol{a}_j$，都存在 $\mathbf{R}^p$ 中的一个向量 $\boldsymbol{c}_j$，使得 $\boldsymbol{a}_j=\boldsymbol{Bc}_j$.

b. 令 $C=[c_1 \quad \cdots \quad c_q]$. 解释为什么存在一个非零向量 u 使得 $Cu=0$.

c. 利用 B 和 C 证明 $Au=0$. 这就说明 A 的列线性相关.

28. 利用习题27证明：如果 $\mathcal{A}$ 和 $\mathcal{B}$ 都是 $\mathbf{R}^n$ 子空间 W 的基，则 $\mathcal{A}$ 中向量的数目不可能多于 $\mathcal{B}$，反过来，$\mathcal{B}$ 中向量的数目也不可能多于 $\mathcal{A}$.

29. [M]设 $H=\text{Span}\{v_1,v_2\}$且 $\mathcal{B}=\{v_1,v_2\}$. 证明 $x\in H$，并求 x 的 $\mathcal{B}$－坐标向量. 其中

$$v_1=\begin{bmatrix}11\\-5\\10\\7\end{bmatrix},\quad v_2=\begin{bmatrix}14\\-8\\13\\10\end{bmatrix},\quad x=\begin{bmatrix}19\\-13\\18\\15\end{bmatrix}$$

30. [M]设 $H=\text{Span}\{v_1,v_2,v_3\}$且 $\mathcal{B}=\{v_1,v_2,v_3\}$. 证明 $\mathcal{B}$ 是 H 的基且 $x\in H$，求 x 的 $\mathcal{B}$－坐标向量. 其中

$$v_1=\begin{bmatrix}-6\\4\\-9\\4\end{bmatrix},\quad v_2=\begin{bmatrix}8\\-3\\7\\-3\end{bmatrix},v_3=\begin{bmatrix}-9\\5\\-8\\3\end{bmatrix},\quad x=\begin{bmatrix}4\\7\\-8\\3\end{bmatrix}$$

基础练习答案

1. 构造 $A=[v_1 \quad v_2 \quad v_3]$，于是由 v_1，v_2，v_3张成的子空间就是 A 的列空间. 这个空间的一组基由 A 的主元列给出.

$$A=\begin{bmatrix}2&3&-1\\-8&-7&6\\6&-1&-7\end{bmatrix}\sim\begin{bmatrix}2&3&-1\\0&5&2\\0&-10&-4\end{bmatrix}\sim\begin{bmatrix}2&3&-1\\0&5&2\\0&0&0\end{bmatrix}$$

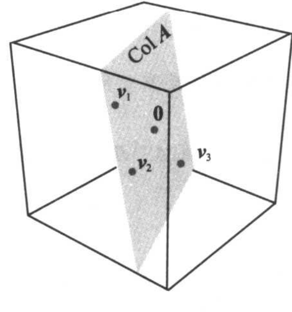
A 的前两列是主元列，它们构成 H 的基. 因此 $\dim H=2$.

182

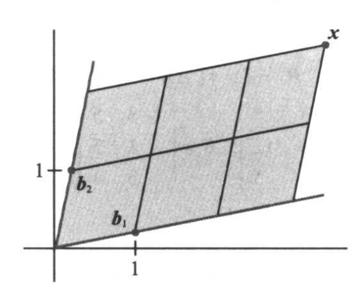
2. 若$[x]_{\mathcal{B}}=\begin{bmatrix}3\\2\end{bmatrix}$，则 x 是以3、2为系数的基向量的线性组合：

$$x=3b_1+2b_2=3\begin{bmatrix}1\\0.2\end{bmatrix}+2\begin{bmatrix}0.2\\1\end{bmatrix}=\begin{bmatrix}3.4\\2.6\end{bmatrix}$$

基$\{b_1,\ b_2\}$确定了 $\mathbf{R}^2$的一个坐标系，如右图网格所示. 注意 x 在 b_1方向上是3个单位，在 b_2方向上是两个单位.



第1题图



第2题图

3. 4维子空间中有由4个线性无关向量构成的基. 这在 $\mathbf{R}^3$ 中是不可能的. 因为 $\mathbf{R}^3$ 中任意线性无关集不超过3个向量，$\mathbf{R}^3$的任意子空间维数不超过3. $\mathbf{R}^3$本身是 $\mathbf{R}^3$空间中唯一的3维子空间. $\mathbf{R}^3$的其他子空间维数为2、1或者0.

第2章补充题

1. 假定下列命题中的矩阵有适当的维度. 判断各个命题的真假，并说明理由.
   a. 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是 $m\times n$ 矩阵，则 $\boldsymbol{AB}^{\mathrm{T}}$ 和 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}$ 都有意义.
   b. 若 $\boldsymbol{AB}=\boldsymbol{C}$ 且 $\boldsymbol{C}$ 有两列，则 $\boldsymbol{A}$ 也有两列.
   c. 用一个对角线上含非零元素的对角阵 $\boldsymbol{A}$ 左乘矩阵 $\boldsymbol{B}$，相当于数乘 $\boldsymbol{B}$ 的行.
   d. 若 $\boldsymbol{BC}=\boldsymbol{BD}$，则 $\boldsymbol{C}=\boldsymbol{D}$.
   e. 若 $\boldsymbol{AC}=\boldsymbol{0}$，则 $\boldsymbol{A}=\boldsymbol{0}$ 或 $\boldsymbol{C}=\boldsymbol{0}$.
   f. 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是 $n\times n$ 矩阵，则 $(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{A}-\boldsymbol{B})=\boldsymbol{A}^2-\boldsymbol{B}^2$.
   g. 一个 $n\times n$ 初等矩阵有 n 个或者 $n+1$ 个非零元素.
   h. 初等矩阵的转置是初等矩阵.
   i. 初等矩阵一定是方阵.
   j. 每个方阵都是一些初等矩阵的乘积.
   k. 若 $\boldsymbol{A}$ 是有3个主元位置的 3×3 矩阵，则存在初等矩阵 $\boldsymbol{E}_1,\cdots,\boldsymbol{E}_p$，使 $\boldsymbol{E}_p\cdots\boldsymbol{E}_1\boldsymbol{A}=\boldsymbol{I}$.
   l. 若 $\boldsymbol{AB}=\boldsymbol{I}$，则 $\boldsymbol{A}$ 可逆.
   m. 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是可逆方阵，则 $\boldsymbol{AB}$ 可逆，且 $(\boldsymbol{AB})^{-1}=\boldsymbol{A}^{-1}\boldsymbol{B}^{-1}$.
   n. 若 $\boldsymbol{AB}=\boldsymbol{BA}$ 且 $\boldsymbol{A}$ 可逆，则 $\boldsymbol{A}^{-1}\boldsymbol{B}=\boldsymbol{BA}^{-1}$.
   o. 若 $\boldsymbol{A}$ 可逆且 $r\neq 0$，则 $(r\boldsymbol{A})^{-1}=r\boldsymbol{A}^{-1}$.
   p. 若 $\boldsymbol{A}$ 是 3×3 矩阵且方程 $\boldsymbol{Ax}=\begin{bmatrix}1\\0\\0\end{bmatrix}$ 有唯一解，则 $\boldsymbol{A}$ 可逆.

2. 求矩阵 $\boldsymbol{C}$，其逆为 $\boldsymbol{C}^{-1}=\begin{bmatrix}4&5\\6&7\end{bmatrix}$.

3. 设 $\boldsymbol{A}=\begin{bmatrix}0&0&0\\1&0&0\\0&1&0\end{bmatrix}$. 证明 $\boldsymbol{A}^3=\boldsymbol{0}$. 利用矩阵代数计算乘积 $(\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2)$.

4. 假设存在某个 $n>1$ 有 $\boldsymbol{A}^n=\boldsymbol{0}$，求 $\boldsymbol{I}-\boldsymbol{A}$ 的逆.

5. 假设 $n\times n$ 矩阵 $\boldsymbol{A}$ 满足方程 $\boldsymbol{A}^2-2\boldsymbol{A}+\boldsymbol{I}=0$，证明 $\boldsymbol{A}^3=3\boldsymbol{A}-2\boldsymbol{I}$ 且 $\boldsymbol{A}^4=4\boldsymbol{A}-3\boldsymbol{I}$.

6. 设 $\boldsymbol{A}=\begin{bmatrix}1&0\\0&-1\end{bmatrix}$，$\boldsymbol{B}=\begin{bmatrix}0&1\\1&0\end{bmatrix}$. 它们是量子力学中用于研究电子自旋的Pauli自旋矩阵. 证明：$\boldsymbol{A}^2=\boldsymbol{I}$，$\boldsymbol{B}^2=\boldsymbol{I}$，且 $\boldsymbol{AB}=-\boldsymbol{BA}$. 满足 $\boldsymbol{AB}=-\boldsymbol{BA}$ 的矩阵称作反交换的.

7. 设 $\boldsymbol{A}=\begin{bmatrix}1&3&8\\2&4&11\\1&2&5\end{bmatrix}$，$\boldsymbol{B}=\begin{bmatrix}-3&5\\1&5\\3&4\end{bmatrix}$. 不事先计算 $\boldsymbol{A}^{-1}$，直接计算 $\boldsymbol{A}^{-1}\boldsymbol{B}$. [提示：$\boldsymbol{A}^{-1}\boldsymbol{B}$ 是方程 $\boldsymbol{AX}=\boldsymbol{B}$ 的解.]

183

8. 求矩阵 $\boldsymbol{A}$，使得变换 $\boldsymbol{x}\mapsto\boldsymbol{Ax}$ 将 $\begin{bmatrix}1\\3\end{bmatrix}$ 和 $\begin{bmatrix}2\\7\end{bmatrix}$ 分别映射到 $\begin{bmatrix}1\\1\end{bmatrix}$ 和 $\begin{bmatrix}3\\1\end{bmatrix}$. [提示：写出包含 $\boldsymbol{A}$ 的一个矩阵方程，然后求解 $\boldsymbol{A}$.]

9. 假设 $\boldsymbol{AB}=\begin{bmatrix}5&4\\-2&3\end{bmatrix}$ 且 $\boldsymbol{B}=\begin{bmatrix}7&3\\2&1\end{bmatrix}$. 求 $\boldsymbol{A}$.

10. 假设 A 可逆. 解释 $A^{\mathrm{T}}A$ 为什么也可逆. 然后证明 $A^{-1}=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$.

11. 设 $x_1,\cdots,x_n$ 是固定的数. 下面的矩阵出现在信号处理、纠错编码和多项式插值等应用中，称作范德蒙(Vandermonde)矩阵.

$$V=\begin{bmatrix}1 & x_1 & x_1^2 & \cdots & x_1^{n-1}\\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1}\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1}\end{bmatrix}$$

已知 $\boldsymbol{y}=(y_1,\cdots,y_n)\in\mathbb{R}^n$，假设 $\boldsymbol{c}=(c_0,\cdots,c_{n-1})\in\mathbb{R}^n$ 满足 $V\boldsymbol{c}=\boldsymbol{y}$，并且定义

$$p(t)=c_0+c_1t+c_2t^2+\cdots+c_{n-1}t^{n-1}$$

a. 证明：$p(x_1)=y_1,\cdots,p(x_n)=y_n$. 我们将 $p(t)$ 称作点 $(x_1,y_1),\cdots,(x_n,y_n)$ 的插值多项式，因为 $p(t)$ 的图像经过这些点.

b. 假设 $x_1,\cdots,x_n$ 是不相同的数，证明 V 的列线性无关. [提示：度数至多 $n-1$ 的多项式有多少个零点？]

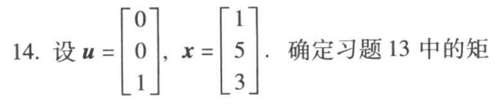
c. 证明："若 $x_1,\cdots,x_n$ 是互不相同的数，$y_1,\cdots,y_n$ 是任意数，存在 $(x_1,y_1),\cdots,(x_n,y_n)$ 的度数 $\leqslant n-1$ 的插值多项式."

12. 设 $A=LU$，其中 L 是可逆下三角矩阵，U 是上三角矩阵. 解释为什么 A 的首列是 L 的首列的倍数. A 的第二列与 L 的列有什么关系呢？

13. 已知 $\boldsymbol{u}\in\mathbb{R}^n$，$\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}=1$，令 $P=\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}$(外积)且 $Q=I-2P$. 验证(a)、(b)和(c).

a. $P^2=P$　　b. $P^{\mathrm{T}}=P$　　c. $Q^2=I$

变换 $\boldsymbol{x}\mapsto P\boldsymbol{x}$ 称作投影，$\boldsymbol{x}\mapsto Q\boldsymbol{x}$ 称作豪斯霍尔德反射. 计算机软件中常利用这类反射生成含多个零元素的向量(通常是矩阵中的一列).



14. 设 $\boldsymbol{u}=\begin{bmatrix}0\\0\\1\end{bmatrix}$，$\boldsymbol{x}=\begin{bmatrix}1\\5\\3\end{bmatrix}$. 确定习题 13 中的矩阵 P 和 Q，并且计算 $P\boldsymbol{x}$ 和 $Q\boldsymbol{x}$. 图 2-29 表明 $Q\boldsymbol{x}$ 和 $\boldsymbol{x}$ 关于 x_1x_2 平面是对称的.

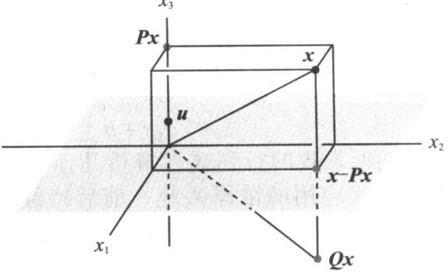


图 2-29　关于平面 $x_3=0$ 的豪斯霍尔德反射

15. 假设 $C=E_3E_2E_1B$，其中 E_1，E_2，E_3 是初等矩阵. 解释 A 为什么与 B 行等价.

16. 设 A 是一个 $n\times n$ 奇异矩阵. 描述如何构造一个非零矩阵 B，使得 $AB=0$.

17. 设 A 是一个 6×4 矩阵，B 是一个 4×6 矩阵. 证明 6×6 矩阵 AB 不可逆.

18. 假设 A 是一个 5×3 矩阵，并且存在一个 3×5 矩阵 C，使得 $CA=I_3$. 另外假定对 $\mathbb{R}^5$ 中某个给定的 $\boldsymbol{b}$，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 至少有一个解. 证明这个解唯一.

19. [M]某些动态系统可以通过观察矩阵的幂来研究，比如在下面这个例子当中. 确定当 k 增加时(例如 $k=2,\cdots,16$)A^k 和 B^k 的变化. 区分 A 和 B 的特点. 进一步考察具有更大解数的同类矩阵，对它们进行猜想.

$$A=\begin{bmatrix}0.4 & 0.2 & 0.3\\ 0.3 & 0.6 & 0.3\\ 0.3 & 0.2 & 0.4\end{bmatrix}\qquad B=\begin{bmatrix}0 & 0.2 & 0.3\\ 0.1 & 0.6 & 0.3\\ 0.9 & 0.2 & 0.4\end{bmatrix}$$

20. [M]设 A_n 是主对角线元素为 0 其余元素为 1 的 $n\times n$ 矩阵. 对 $n=4,5,6$ 计算 A_n^{-1}，对于更大的 n，猜想 A_n^{-1} 的一般形式.

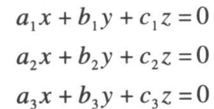
184

第3章　行　列　式

实例介绍：解析几何中的行列式

行列式是由一些数值排列成的方阵经计算得到的一个数. 早在 1683 年和 1693 年，日本数学家关孝和和德国数学家莱布尼兹就分别独立地提出了行列式的概念. 以后很长一段时间内，行列式主要应用于讨论线性方程组. 约 160 年后，行列式发展成为矩阵的一个独立的理论分支.

1750 年，瑞士数学家克莱姆在他的论文中提到行列式或许也可以应用于解析几何中. 在这篇论文中，克莱姆利用行列式构造了 xy 平面中某些曲线的方程. 同时，他在文中提出了利用行列式求解 $n\times n$ 线性方程组的著名法则. 紧接着，1812 年，柯西发表论文，用行列式给出了计算某些实心多面体体积的公式，并且将这些公式与先前行列式的研究结果联系起来. 柯西所讨论的“晶体”包括图 3-1 的四面体和图 3-2 的平行六面体. 如果平行六面体的四个顶点的坐标分别是原点 $\mathbf{0}=(0,0,0)$，$\boldsymbol{v}_1=(a_1,b_1,c_1)$，$\boldsymbol{v}_2=(a_2,b_2,c_2)$和 $\boldsymbol{v}_3=(a_3,b_3,c_3)$，则其体积等于下列方程组系数矩阵行列式的绝对值：



$$a_1x+b_1y+c_1z=0$$
$$a_2x+b_2y+c_2z=0$$
$$a_3x+b_3y+c_3z=0$$

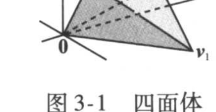


图 3-1　四面体

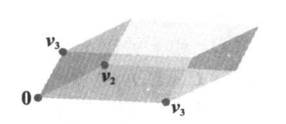


图 3-2　平行六面体

柯西所发现的行列式在解析几何中的应用激起了人们探究行列式应用的浓厚兴趣，前后持续了近 100 年. 托马斯·米尔在他的一部四卷著作中概述了 20 世纪初期以前人们的发现和成果.

在柯西所处的时代，人们讨论的矩阵其维度通常很小，行列式在解析几何以及数学的其他分支中都扮演着很重要的角色. 如今，常见的大型矩阵计算中，行列式的数值意义已经不大. 不过，行列式公式仍然可以给出矩阵的重要信息，而且在线性代数的某些应用中行列式的知识依然很有用.

本章我们有三个目标：一是证明方阵 $\boldsymbol{A}$ 可逆的一个判别法，它与 $\boldsymbol{A}$ 的元素而非 $\boldsymbol{A}$ 的列有关；二是给出 $\boldsymbol{A}^{-1}$ 以及 $\boldsymbol{A}^{-1}\boldsymbol{b}$ 的计算公式，它们在理论应用上十分有用；最后给出本章实例介绍中提到的行列式的几何解释. 第一个目标在 3.2 节中实现，其余两个则放在 3.3 节中解决.

3.1　行列式简介

回忆 2.2 节，我们知道一个 2×2 的矩阵可逆当且仅当其行列式非零. 为了将此

结论推广到维度较大的矩阵上，我们需要给出 $n\times n$ 矩阵行列式的定义. 对于 3×3 的情形，其定义可以通过观察一个 3×3 可逆矩阵的行变换过程来得到.

考虑 $A=[a_{ij}]$，其中 $a_{11}\neq0$. 如果用 a_{11} 分别乘以矩阵 A 的第二行和第三行，然后从第二、三行减去第一行的适当倍数，我们发现 A 与下面两个矩阵行等价：

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{11}a_{21} & a_{11}a_{22} & a_{11}a_{23} \\ a_{11}a_{31} & a_{11}a_{32} & a_{11}a_{33} \end{bmatrix} \sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22}-a_{12}a_{21} & a_{11}a_{23}-a_{13}a_{21} \\ 0 & a_{11}a_{32}-a_{12}a_{31} & a_{11}a_{33}-a_{13}a_{31} \end{bmatrix} \tag{1}$$

由于 A 是可逆的，公式(1)右侧矩阵的(2,2)－元或(3,2)－元非零. 我们假定(2,2)－元非零(否则在进行下面运算之前可以先做一次行交换)，用 $a_{11}a_{22}-a_{12}a_{21}$ 乘以第三行，然后给新得到的第三行加上第二行的 $-(a_{11}a_{32}-a_{12}a_{31})$ 倍. 这就得到：

$$A \sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22}-a_{12}a_{21} & a_{11}a_{23}-a_{13}a_{21} \\ 0 & 0 & a_{11}\Delta \end{bmatrix}$$

其中

$$\Delta = a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31} \tag{2}$$

由于 A 是可逆的，Δ 一定非零. 我们将在 3.2 节中看到反之亦然. 我们把公式(2)中的 Δ 称为 3×3 矩阵 A 的**行列式**(determinant).

回忆 2×2 矩阵 $A=[a_{ij}]$的行列式是这样一个数：

$$\det A = a_{11}a_{22}-a_{12}a_{21}$$

对 1×1 矩阵，即 $A=[a_{11}]$，我们定义 $\det A=a_{11}$. 为了将行列式的定义推广到高阶矩阵，下面我们利用 2×2 行列式来重写 3×3 行列式 Δ. 由于 Δ 中的项可以分组为： 186

$$\Delta = a_{11}\cdot\det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12}\cdot\det\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13}\cdot\det\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

简单起见，我们记：

$$\Delta = a_{11}\cdot\det A_{11} - a_{12}.\det A_{12} + a_{13}\cdot\det A_{13} \tag{3}$$

其中 A_{11}，A_{12} 以及 A_{13} 都是从 A 中划去第一行以及三列中的一列而得到的. 对于任意方阵 A，用 A_{ij} 表示划去矩阵 A 中第 i 行和第 j 列以后得到的子矩阵. 例如，如果

$$A = \begin{bmatrix} 1 & -2 & 5 & 0 \\ 2 & 0 & 4 & -1 \\ 3 & 1 & 0 & 7 \\ 0 & 4 & -2 & 0 \end{bmatrix}$$

则 A_{32} 就是划去 A 中第三行和第二列得到的. 即

$$\begin{bmatrix} 1 & -2 & 5 & 0 \\ 2 & 0 & 4 & -1 \\ 3 & 1 & 0 & 7 \\ 0 & 4 & -2 & 0 \end{bmatrix}$$

于是,

$$A_{32}=\begin{bmatrix}1&5&0\\2&4&-1\\0&-2&0\end{bmatrix}$$

至此，我们可以给出行列式的一种递归定义. 当 $n=3$ 时，$\det A$ 可以由 2×2 阶子矩阵 A_{1j} 的行列式来定义，如公式(3)所示. 当 $n=4$ 时，$\det A$ 同样可以利用 3×3 阶子矩阵 A_{1j} 的行列式来定义. 一般的，一个 $n\times n$ 矩阵的行列式可以由 $(n-1)\times(n-1)$ 阶子矩阵的行列式来定义.

【定义】　当 $n\geqslant2$，一个 $n\times n$ 矩阵 $A=[a_{ij}]$ 的**行列式**(determinant)是形如 $\pm a_{1j}\det A_{1j}$ 的 n 项之和，其中各项正负交替，$a_{11},a_{12},\cdots,a_{1n}$ 是矩阵 A 第一行的元素. 用符号表示为:

$$\det A=a_{11}\det A_{11}-a_{12}\det A_{12}+\cdots+(-1)^{1+n}a_{1n}\det A_{1n}=\sum_{j=1}^{n}(-1)^{1+j}a_{1j}\det A_{1j}$$

【例题1】　计算下列矩阵 A 的行列式

$$A=\begin{bmatrix}1&5&0\\2&4&-1\\0&-2&0\end{bmatrix}$$

187

解: 首先计算 $\det A=a_{11}\det A_{11}-a_{12}\det A_{12}+a_{13}\det A_{13}$:

$$\det A=1\cdot\det\begin{bmatrix}4&-1\\-2&0\end{bmatrix}-5\cdot\det\begin{bmatrix}2&-1\\0&0\end{bmatrix}+0\cdot\det\begin{bmatrix}2&4\\0&-2\end{bmatrix}$$

$$=1(0-2)-5(0-0)+0(-4-0)=-2$$ ■

另一种矩阵行列式的常见记法是用一对竖线代替方括号. 上面的计算过程可以写成:

$$\det A=1\begin{vmatrix}4&-1\\-2&0\end{vmatrix}-5\begin{vmatrix}2&-1\\0&0\end{vmatrix}+0\begin{vmatrix}2&4\\0&-2\end{vmatrix}=\cdots=-2$$

为了引出下一个定理，我们将用一种与上文略有差异的形式定义 $\det A$. 已知 $A=[a_{ij}]$，A 的 (i,j) - **余子式**((i,j) - cofactor)是数值 C_{ij}，它由下式给出:

$$C_{ij}=(-1)^{i+j}\det A_{ij} \tag{4}$$

于是,

$$\det A=a_{11}C_{11}+a_{12}C_{12}+\cdots+a_{1n}C_{1n}$$

这个公式被称作 **A 按第一行的余子式展开**(cofactor expansion across the first row). 以下基本定理的证明较长，在此略过.

【定理1】　$n\times n$ 矩阵 A 的行列式可以根据按任意行或列的余子式展开进行计算. 使用(4)中的余子式，按第 i 行的余子式展开是:

$$\det A=a_{i1}C_{i1}+a_{i2}C_{i2}+\cdots+a_{in}C_{in}$$

按第 j 列的余子式展开是:

$$\det A=a_{1j}C_{1j}+a_{2j}C_{2j}+\cdots+a_{nj}C_{nj}$$

其中 (i,j) - 余子式的正负号取决于 a_{ij} 在原矩阵 A 中的位置，与 a_{ij} 本身的符号无

关. 因子$(-1)^{i+j}$所确定的矩阵各个位置上的符号呈现如下棋盘图案：

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \\ + & - & + & \\ \vdots & & & \ddots \end{bmatrix}$$

【例题 2】 按第三行的余子式展开，计算 det $\boldsymbol{A}$，其中

$$\boldsymbol{A}=\begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

188

解：计算

$$\begin{aligned} \det \boldsymbol{A} &= a_{31}C_{31}+a_{32}C_{32}+a_{33}C_{33} \\ &= (-1)^{3+1}a_{31}\det A_{31}+(-1)^{3+2}a_{32}\det A_{32}+(-1)^{3+3}a_{33}\det A_{33} \\ &= 0\begin{vmatrix} 5 & 0 \\ 4 & -1 \end{vmatrix} - (-2)\begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} + 0\begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix} \\ &= 0+2(-1)+0=-2 \end{aligned}$$

■

定理 1 对计算含多个零元素的矩阵的行列式很有用. 例如，假定矩阵中有一行元素大部分为零，则按该行的余子式展开式中也会有多项为零，这就意味着有很多项余子式不必计算. 对于某一列含多个零元素的情形，这个方法也适用.

【例题 3】 计算 det $\boldsymbol{A}$，其中

$$\boldsymbol{A}=\begin{bmatrix} 3 & -7 & 8 & 9 & -6 \\ 0 & 2 & -5 & 7 & 3 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 2 & 4 & -1 \\ 0 & 0 & 0 & -2 & 0 \end{bmatrix}$$

解：矩阵 $\boldsymbol{A}$ 按第一列的余子式展开式中，除第一项外，其余都等于零. 因此：

$$\det \boldsymbol{A}=3\cdot\begin{vmatrix} 2 & -5 & 7 & 3 \\ 0 & 1 & 5 & 0 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & -2 & 0 \end{vmatrix}-0\cdot C_{21}+0\cdot C_{31}-0\cdot C_{41}+0\cdot C_{51}$$

以后我们将略去余子式展开式中的零项. 接下来，对这个 4×4 矩阵按第一列展开，此处仍然利用该列大部分元素为零的特点，有：

$$\det \boldsymbol{A}=3\cdot2\cdot\begin{vmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{vmatrix}$$

这个 3×3 行列式就是我们在例题 1 中计算过的，其结果为 -2. 因此，$\det \boldsymbol{A}=3\times2\times(-2)=-12$. ■

例题 3 中的矩阵近似于三角阵. 我们可以对其所用的方法加以改造，完成下述定理的证明.

189

【定理2】 若 A 是三角阵，则 $\det \boldsymbol{A}$ 是 $\boldsymbol{A}$ 的主对角线元素的乘积.

当矩阵中某一行或某一列元素全为零时，例题3中寻找零元素的策略发挥到了极至. 在这种情况下，按这样一行或一列的余子式展开是一些零的和，因此该矩阵的行列式为零. 然而，大多数余子式展开都不这么容易计算.

注记

按照现在的标准，一个 25×25 矩阵只是一个小矩阵. 然而我们不可能通过余子式展开的方法来计算 25×25 矩阵的行列式. 一般地，计算一个余子式展开需要超过 $n!$ 次乘法运算，而 $25!$ 约等于 1.5×10^{25}.

假定计算机每秒钟可以执行一万亿次乘法运算，那么计算机采用这种方法计算 25×25 阶矩阵的行列式需要超过 500 000 年的时间. 庆幸的是，我们很快就能找出快捷的方法来解决这个问题.

习题19～38主要讨论 2×2 的情形，揭示了这类行列式的重要性质. 下一节将利用习题33～36的结果推导出 $n\times n$ 矩阵的类似性质.

基础练习

计算 $\begin{vmatrix} 5 & -7 & 2 & 2 \\ 0 & 3 & 0 & -4 \\ -5 & -8 & 0 & 3 \\ 0 & 5 & 0 & -6 \end{vmatrix}$.

习题 3.1

按第一行进行余子式展开，计算习题1～8中各矩阵的行列式. 另外对于习题1～4，按第二列的余子式展开计算行列式.

1. $\begin{vmatrix} 3 & 0 & 4 \\ 2 & 3 & 2 \\ 0 & 5 & -1 \end{vmatrix}$

2. $\begin{vmatrix} 0 & 5 & 1 \\ 4 & -3 & 0 \\ 2 & 4 & 1 \end{vmatrix}$

3. $\begin{vmatrix} 2 & -4 & 3 \\ 3 & 1 & 2 \\ 1 & 4 & -1 \end{vmatrix}$

4. $\begin{vmatrix} 1 & 3 & 5 \\ 2 & 1 & 1 \\ 3 & 4 & 2 \end{vmatrix}$

5. $\begin{vmatrix} 2 & 3 & -4 \\ 4 & 0 & 5 \\ 5 & 1 & 6 \end{vmatrix}$

6. $\begin{vmatrix} 5 & -2 & 4 \\ 0 & 3 & -5 \\ 2 & -4 & 7 \end{vmatrix}$

7. $\begin{vmatrix} 4 & 3 & 0 \\ 6 & 5 & 2 \\ 9 & 7 & 3 \end{vmatrix}$

8. $\begin{vmatrix} 8 & 1 & 6 \\ 4 & 0 & 3 \\ 3 & -2 & 5 \end{vmatrix}$

利用余子式展开计算习题9～14中各矩阵的行列式，每一步中，选择计算量最小的行或列进行展开.

9. $\begin{vmatrix} 6 & 0 & 0 & 5 \\ 1 & 7 & 2 & -5 \\ 2 & 0 & 0 & 0 \\ 8 & 3 & 1 & 8 \end{vmatrix}$

10. $\begin{vmatrix} 1 & -2 & 5 & 2 \\ 0 & 0 & 3 & 0 \\ 2 & -6 & -7 & 5 \\ 5 & 0 & 4 & 4 \end{vmatrix}$

11. $\begin{vmatrix} 3 & 5 & -8 & 4 \\ 0 & -2 & 3 & -7 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 2 \end{vmatrix}$

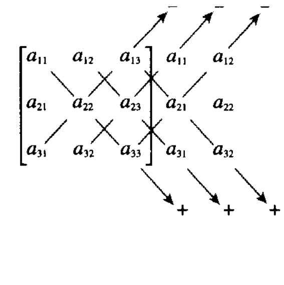
12. $\begin{vmatrix} 4 & 0 & 0 & 0 \\ 7 & -1 & 0 & 0 \\ 2 & 6 & 3 & 0 \\ 5 & -8 & 4 & -3 \end{vmatrix}$

13. $\begin{vmatrix} 4 & 0 & -7 & 3 & -5 \\ 0 & 0 & 2 & 0 & 0 \\ 7 & 3 & -6 & 4 & -8 \\ 5 & 0 & 5 & 2 & -3 \\ 0 & 0 & 9 & -1 & 2 \end{vmatrix}$

14. $\begin{vmatrix} 6 & 3 & 2 & 4 & 0 \\ 9 & 0 & -4 & 1 & 0 \\ 8 & -5 & 6 & 7 & 1 \\ 3 & 0 & 0 & 0 & 0 \\ 4 & 2 & 3 & 2 & 0 \end{vmatrix}$

190

3×3 行列式的展开式可按右图记忆. 先将原矩阵的前两列照抄在矩阵的右边，然后根据六条对角线上元素的乘积计算行列式. 累加方向向下的对角线上元素的乘积，再减去方向向上的对角线上元素的乘积，即可求出行列式. 使用上述方法计算习题 15 ~ 18的行列式. 警告：此方法不适用于 4×4 或更高阶的矩阵.



15. $\begin{vmatrix} 3 & 0 & 4 \\ 2 & 3 & 2 \\ 0 & 5 & -1 \end{vmatrix}$　　16. $\begin{vmatrix} 0 & 5 & 1 \\ 4 & -3 & 0 \\ 2 & 4 & 1 \end{vmatrix}$

17. $\begin{vmatrix} 2 & -4 & 3 \\ 3 & 1 & 2 \\ 1 & 4 & -1 \end{vmatrix}$　　18. $\begin{vmatrix} 1 & 3 & 5 \\ 2 & 1 & 1 \\ 3 & 4 & 2 \end{vmatrix}$

分析习题 19 ~ 24 中初等行变换对矩阵行列式的影响. 对每一道题，指出行变换并说明它是如何影响行列式的.

19. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} c & d \\ a & b \end{bmatrix}$　　20. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a & b \\ kc & kd \end{bmatrix}$

21. $\begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}, \begin{bmatrix} 3 & 4 \\ 5+3k & 6+4k \end{bmatrix}$　　22. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a+kc & b+kd \\ c & d \end{bmatrix}$

23. $\begin{bmatrix} 1 & 1 & 1 \\ -3 & 8 & -4 \\ 2 & -3 & 2 \end{bmatrix}, \begin{bmatrix} k & k & k \\ -3 & 8 & -4 \\ 2 & -3 & 2 \end{bmatrix}$　　24. $\begin{bmatrix} a & b & c \\ 3 & 2 & 2 \\ 6 & 5 & 6 \end{bmatrix}, \begin{bmatrix} 3 & 2 & 2 \\ a & b & c \\ 6 & 5 & 6 \end{bmatrix}$

计算习题 25 ~ 30 中初等矩阵的行列式(见 2.2 节).

25. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{bmatrix}$　　26. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{bmatrix}$　　27. $\begin{bmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

28. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}$　　29. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$　　30. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

利用习题 25 ~ 28 回答习题 31 和 32.

31. 初等行替换矩阵的行列式是多少?

32. 数量 k 所对应的初等数乘矩阵的行列式是多少?

验证习题 33 ~ 36 的矩阵满足公式 $\det EA = (\det E)(\det A)$. 其中 E 是所给初等矩阵，

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

33. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$　　34. $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$　　35. $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$　　36. $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$

37. 设 $A = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}$，写出 $5A$. 是否有 $\det 5A = 5\det A$?

38. 设 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$，设 k 是一个数量. 求 $\det kA$ 与 k，$\det A$ 的关联公式.

在习题39和40中，设A是$n\times n$矩阵. 判断下列命题的真假，并说明理由.

39. a. $n\times n$矩阵的行列式由其$(n-1)\times(n-1)$子矩阵的行列式来定义.
    b. A的(i,j)-余子式是从A中划去第i行和第j列后得到的矩阵A_{ij}.
40. a. $\det A$按某一列的余子式展开等于按某一行的余子式展开乘以-1.
    b. 三角矩阵的行列式是其主对角线元素之和.
41. 设$u=\begin{bmatrix}3\\0\end{bmatrix}$，$v=\begin{bmatrix}1\\2\end{bmatrix}$. 计算由$u$，$v$，$u+v$和$0$所确定的平行四边形的面积以及$[u\quad v]$的行列式. 两者相比结果如何？用随机数$x$代替$v$的第一个元素，重复上述过程. 画图并且解释你的发现.

191

42. 设$u=\begin{bmatrix}a\\b\end{bmatrix}$，$v=\begin{bmatrix}c\\0\end{bmatrix}$，其中$a$，$b$，$c$均为正数(简单起见). 计算由$u$，$v$，$u+v$和$0$所确定的平行四边形的面积以及$[u\quad v]$和$[v\quad u]$的行列式. 画图并且解释你的发现.
43. [M] $\det(A+B)=\det A+\det B$是否成立？生成5×5随机矩阵A，B并计算$\det(A+B)-\det A-\det B$，以检验上述猜想. (参考2.1节习题37.)对于不同的n，另选三对$n\times n$矩阵重复上述计算，并报告你的结论(参考2.1节习题37).
44. [M] $\det(AB)=(\det A)(\det B)$是否成立？仿照习题43，用4对随机矩阵进行检验，然后做出猜想.
45. [M] 构造4×4随机矩阵A，其元素均为-9到9之间的整数. 比较$\det A$和$\det A^{T}$，$\det(-A)$，$\det(2A)$，$\det(10A)$的大小. 另选两个4×4随机整数矩阵重复上述过程，然后猜想这些行列式之间的关系. 用若干5×5、6×6随机整数矩阵检验你的猜想，如有必要可以进行修正，最后报告你的结论(参考2.1节习题36).
46. [M] $\det A$和$\det A^{-1}$有什么关系？对$n=4$，5和6，用$n\times n$随机整数矩阵来试验，并做出猜想. *注意*：如果矩阵的行列式恰好为0，先将矩阵变换为阶梯形式，再进行讨论.

基础练习答案

充分利用矩阵中的零元素. 先按第三列进行余子式展开，得到一个3×3矩阵，再按其第一列进行余子式展开即可.

$$\begin{vmatrix}5&-7&2&2\\0&3&0&-4\\-5&-8&0&3\\0&5&0&-6\end{vmatrix}=(-1)^{1+3}2\begin{vmatrix}0&3&-4\\-5&-8&3\\0&5&-6\end{vmatrix}=2\cdot(-1)^{2+1}(-5)\begin{vmatrix}3&-4\\5&-6\end{vmatrix}=20$$

其中倒数第二个等号后的$(-1)^{2+1}$是由于-5在3×3矩阵中处在$(2,1)$的位置.

3.2 行列式的性质

行列式的奥妙在于对矩阵进行了行变换之后，行列式会发生相应的变化. 下面的定理推广了3.1节习题19~24的结论，其证明将在本节末尾给出.

【定理3】 行变换

设A是一个方阵.

a. 如果将A的某一行乘以某数加到另一行得到矩阵B，则$\det A=\det B$.

b. 如果交换A的两行得到矩阵B，则$\det B=-\det A$.

c. 如果将A的某一行数乘以k得到矩阵B，则$\det B=k\cdot\det A$.

192

下面的例题演示了如何利用定理3高效地计算行列式.

【例题1】 计算 $\det A$，其中 $A=\begin{bmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{bmatrix}$.

解： 解题的方法是将 A 化简为阶梯形式，然后利用"三角矩阵的行列式等于其对角线元素之积"这个事实. 为将第1列下方的元素化为零，我们做两次行替换，它们不改变行列式的值，即

$$\det A=\begin{vmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{vmatrix}=\begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ -1 & 7 & 0 \end{vmatrix}=\begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{vmatrix}$$

交换第2行和第3行，行列式变号，因此，

$$\det A=-\begin{vmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{vmatrix}=-(1)(3)(-5)=15$$ ■

定理3(c)多用于提取矩阵中某行的公共因子，比如：

$$\begin{vmatrix} * & * & * \\ 5k & -2k & 3k \\ * & * & * \end{vmatrix}=k\begin{vmatrix} * & * & * \\ 5 & -2 & 3 \\ * & * & * \end{vmatrix}$$

其中以星号标记的元素不变，在下面的例题中我们将用到这个性质.

【例题2】 计算 $\det A$，其中 $A=\begin{bmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{bmatrix}$.

解： 为了简化计算，我们希望将左上角的元素化为1. 通过交换第1行和第4行可以做到这一点，不过这里采用另一种方法. 首先提取第一行的因子2，然后对矩阵做行替换，以得到第1列下方的零：

$$\det A=2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{vmatrix}=2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & -12 & 10 & 10 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

接下来，可以提取第3行的因子2，也可以把第2列的"3"作为主元. 这里选择后一种操作，将第2行的4倍加到第3行：

$$\det A=2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

193

最后，将第3行的 $-1/2$ 倍加到第4行，然后计算"三角矩阵"的行列式：

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 2\cdot(1)(3)(-6)(1) = -36$$ ■

假设方阵 A 通过行替换和行交换被化简为阶梯形式 U(这通常是可以做到的，参考1.2节的行化简算法). 如果进行了 r 次行交换，则根据定理3应该有：

$$\det A = (-1)^r \det U$$

由于 U 是阶梯形式，U 是一个三角矩阵，因此 det U 是对角线元素 $u_{11},\cdots,u_{nn}$ 的乘积. 如果 A 可逆，则元素 u_{ii} 都是主元(因为 $A\sim I_n$ 且 u_{ii} 未被缩放到1)，否则，至少 u_{nn} 是0，因此乘积 $u_{11}\cdots u_{nn}$ 等于0（见图3-3).

$$U=\begin{bmatrix} ■ & * & * & * \\ 0 & ■ & * & * \\ 0 & 0 & ■ & * \\ 0 & 0 & 0 & ■ \end{bmatrix} \quad U=\begin{bmatrix} ■ & * & * & * \\ 0 & ■ & * & * \\ 0 & 0 & 0 & ■ \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

det $U\neq 0$　　　　det $U = 0$

图3-3　典型的阶梯形矩阵

这样就有，

$$\det A = \begin{cases} (-1)^r\cdot(U\text{ 中主元之积}) & \text{当 } A \text{ 可逆时} \\ 0 & \text{当 } A \text{ 不可逆时} \end{cases} \tag{1}$$

有趣的是，尽管阶梯形式 U 不唯一(因为行化简并不完全)，主元也不唯一，但是除符号变化外，主元的乘积是唯一的.

公式(1)不仅给了计算 det A 的具体解释，而且证明了本节最主要的定理：

【定理4】　方阵 A 可逆当且仅当 det $A\neq 0$.

定理4为可逆矩阵定理补充了一个新的等价命题“det $A\neq 0$”. 它的一个有用的推论是，如果矩阵 A 的各列线性相关，则 det $A=0$. 此外，如果矩阵 A 的各行线性相关，则 det $A=0$(矩阵 A 的行是其转置矩阵 A^T 的列，转置矩阵 A^T 的各列线性相关表明 A^T 是奇异矩阵. 由可逆矩阵定理可知，若 A^T 是奇异矩阵，则 A 也是奇异矩阵). 实际上，当矩阵中包含相同的两行或两列，或者包含一个零行或零列，则线性相关是显然的.

194

【例题3】　计算 det A，其中 $A=\begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ -6 & 7 & -7 & 4 \\ -5 & -8 & 0 & 9 \end{bmatrix}$.

解：将第1行的2倍加到第3行，由于所得的矩阵中第2、3行相同，我们有

$$\det A = \det\begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 5 & -3 & -6 \\ -5 & -8 & 0 & 9 \end{bmatrix} = 0$$ ■

注记

1. 大部分用于计算普通矩阵行列式 $\det A$ 的计算机程序都使用上面的公式(1).

2. 可以证明，利用行变换计算 $n\times n$ 矩阵的行列式需要大约 $2n^3/3$ 次算术运算. 任何一台现代微型计算机都可以在几分之一秒内计算出 25×25 矩阵的行列式，运算量只有 10 000 次左右.

利用大型“稀疏”矩阵中存在大量零元素这一特点，计算机也可以十分方便地处理这类矩阵. 当然，矩阵中的零元素也可以提高手算的速度. 下面例题的计算过程结合使用了行变换以及 3.1 节所介绍的利用零元素进行余子式展开的策略.

【例题 4】 计算 $\det A$，其中 $A=\begin{bmatrix}0&1&2&-1\\2&5&-7&3\\0&3&6&2\\-2&-5&4&-2\end{bmatrix}$.

解： 解此题一个很好的方法是将第 1 列中的 2 作为主元，消除其下方的 -2. 然后利用余子式展开降低行列式的阶数，再做一次行替换，得到

$$\det A=\begin{vmatrix}0&1&2&-1\\2&5&-7&3\\0&3&6&2\\0&0&-3&1\end{vmatrix}=-2\begin{vmatrix}1&2&-1\\3&6&2\\0&-3&1\end{vmatrix}=-2\begin{vmatrix}1&2&-1\\0&0&5\\0&-3&1\end{vmatrix}$$

交换第 2、3 行将得出一个“三角行列式”. 另一种方法是按第一列进行余子式展开：

$$\det A=(-2)(1)\begin{vmatrix}0&5\\-3&1\end{vmatrix}=-2\cdot(15)=-30$$ ■ 195

3.2.1 列变换

与前面讨论的行变换类似，我们也可以对矩阵施行列变换. 下面的定理表明，列变换与行变换对于行列式的变化具有相同的效果.

【定理 5】 若 A 是 $n\times n$ 矩阵，则 $\det A^T=\det A$.

证明： $n=1$ 时显然. 假设该定理对于 $k\times k$ 行列式成立，考虑 $n=k+1$ 的情况. 由于 A 中 a_{1j} 的余子式与 A^T 中 a_{j1} 的余子式都是 $k\times k$ 行列式，根据归纳假设可知两者相等. 因此 $\det A$ 按第一行的余子式展开等于 $\det A^T$ 按第一列的余子式展开，即 A 与 A^T 的行列式相等. 这样，定理对 $n=k+1$ 成立，并且若对某个 n 成立则对 $n+1$ 亦成立. 根据归纳法，定理对一切 $n\geq 1$ 都成立. ■

根据定理 5，将定理 3 各命题中的“行”换为“列”，命题依然成立. 为验证这一性质，只需对 A^T 应用原来的定理 3，A^T 上的行变换实际相当于 A 上的列变换.

列变换有一定的理论意义，手算时也十分有用. 不过为简单起见，我们在数值计算中只施行行变换.

3.2.2 行列式与矩阵乘积

下述定理的证明在本节末尾给出，其应用见习题部分.

【定理 6】 乘法性质

若 A 和 B 都是 $n\times n$ 矩阵，则 $\det AB=(\det A)(\det B)$.

【例题5】　对 $A=\begin{bmatrix}6 & 1\\3 & 2\end{bmatrix}$ 和 $B=\begin{bmatrix}4 & 3\\1 & 2\end{bmatrix}$ 验证定理6.

解：

$$AB=\begin{bmatrix}6 & 1\\3 & 2\end{bmatrix}\begin{bmatrix}4 & 3\\1 & 2\end{bmatrix}=\begin{bmatrix}25 & 20\\14 & 13\end{bmatrix}$$

且

$$\det AB = 25\cdot 13-20\cdot 14=325-280=45$$

因为 $\det A=9$，$\det B=5$，我们有

196

$$(\det A)(\det B)=9\cdot 5=45=\det AB$$ ■

警告：有一种普遍存在的误解，认为对于矩阵的和也有类似于定理6的结论. 但是，$\det(A+B)$ 一般不等于 $\det A+\det B$.

3.2.3　行列式函数的线性性质

对于 $n\times n$ 矩阵 A，我们可以将 $\det A$ 看作是 A 中 n 个列向量的函数. 以下将证明：如果 A 中除某列以外其余列均固定，则 $\det A$ 就是允许变化的那一列的一个线性函数.

假设允许 A 中第 j 列有变化，并且记

$$A=[\boldsymbol{a}_1\ \cdots\ \boldsymbol{a}_{j-1}\ \ \boldsymbol{x}\ \ \boldsymbol{a}_{j+1}\ \cdots\ \boldsymbol{a}_n]$$

定义从 $\mathbf{R}^n$ 到 $\mathbf{R}$ 的变换 T 为

$$T(\boldsymbol{x})=\det[\boldsymbol{a}_1\ \cdots\ \boldsymbol{a}_{j-1}\ \ \boldsymbol{x}\ \ \boldsymbol{a}_{j+1}\ \cdots\ \boldsymbol{a}_n]$$

则

$$T(c\boldsymbol{x})=cT(\boldsymbol{x}),\text{对任意数量 } c \text{ 及任意 } \boldsymbol{x}\in\mathbf{R}^n \tag{2}$$

$$T(\boldsymbol{u}+\boldsymbol{v})=T(\boldsymbol{u})+T(\boldsymbol{v}),\text{对任意 } \boldsymbol{u},\boldsymbol{v}\in\mathbf{R}^n \tag{3}$$

性质(2)是定理3(c)应用于 A 中列的结果. 性质(3)的证明可以利用 $\det A$ 按第 j 列的余子式展开来完成(见习题43)这种行列式线性性质在高等课程中有许多有用的推论.

3.2.4　定理3和定理6的证明

使用2.2节中介绍的初等矩阵来叙述定理3，证明起来将十分方便. 设 E 为一个初等矩阵，若 E 是将单位矩阵某一行的倍数加到另一行所得的矩阵，则称 E 为行替换(矩阵)；若 E 是交换 I 的两行所得的矩阵，则 E 为行交换；若 E 是用非零数量 r 乘以 I 中某行所得的矩阵，则称 E 为 r 数乘. 采用这种定义，定理3可以重写成：

若 A 是一个 $n\times n$ 矩阵，且 E 是 $n\times n$ 初等矩阵，则

$$\det EA=(\det E)(\det A)$$

其中，

$$\det E=\begin{cases}1 & \text{若 } E \text{ 是一个行替换}\\-1 & \text{若 } E \text{ 是一个交换}\\r & \text{若 } E \text{ 是一个 } r \text{ 数乘}\end{cases}$$

定理3的证明： 对 A 的维度进行归纳证明. 2×2 矩阵的情形已在3.1章的习题33～36中验证过. 假设定理对于 $k\times k$ 矩阵的行列式成立，其中 $k\geqslant 2$. 令 $n=k+1$，A 是

一个 $n\times n$ 矩阵. $\boldsymbol{E}$ 在 $\boldsymbol{A}$ 上的作用只能改变 $\boldsymbol{A}$ 的一行或者两行，因此我们可以假设 $\boldsymbol{A}$ 的第 i 行在 $\boldsymbol{E}$ 作用下不变，并且按这行展开 $\det \boldsymbol{EA}$. 令 $\boldsymbol{A}_{ij}$ 是从 $\boldsymbol{A}$ 中划去第 i 行、第 j 列所得的矩阵，$\boldsymbol{B}_{ij}$ 是从 $\boldsymbol{EA}$ 中划去第 i 行、第 j 列所得的矩阵. 则 $\boldsymbol{B}_{ij}$ 的行可以由 $\boldsymbol{A}_{ij}$ 的行经过 $\boldsymbol{A}$ 上的某个初等行变换得到，且这个初等行变换的类型与 $\boldsymbol{E}$ 相同. 因为这些子矩阵都是 $k\times k$ 矩阵，根据归纳假设，有：

[197]

$$\det \boldsymbol{B}_{ij}=\alpha\cdot\det \boldsymbol{A}_{ij}$$

其中 $\alpha=1$, -1 或 r，其值取决于 $\boldsymbol{E}$ 的类型. 按第 i 行进行余子式展开，得

$$\begin{aligned}\det \boldsymbol{EA}&=a_{i1}(-1)^{i+1}\det \boldsymbol{B}_{i1}+\cdots+a_{in}(-1)^{i+n}\det \boldsymbol{B}_{in}\\&=\alpha a_{i1}(-1)^{i+1}\det \boldsymbol{A}_{i1}+\cdots+\alpha a_{in}(-1)^{i+n}\det \boldsymbol{A}_{in}\\&=\alpha\cdot\det \boldsymbol{A}\end{aligned}$$

特别地，令 $\boldsymbol{A}=\boldsymbol{I}_n$，我们可以看到，对于不同类型的 $\boldsymbol{E}$，$\det \boldsymbol{E}=-1$，1 或 r. 于是该定理对 $n=k+1$ 成立，并且，如果定理对某个 n 成立，则对 $n+1$ 亦成立. 根据归纳法，定理对一切 $n\geq 2$ 都成立. 定理在 $n=1$ 时成立则是显见的. ■

定理 6 的证明： 由 2.3 节习题 27 知，若 $\boldsymbol{A}$ 不可逆，则 $\boldsymbol{AB}$ 也不可逆. 此时有 $\det \boldsymbol{AB}=(\det \boldsymbol{A})(\det \boldsymbol{B})$，因为根据定理 4 等式两端的值均为零. 若 $\boldsymbol{A}$ 可逆，则由可逆矩阵定理可知，矩阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{I}_n$ 行等价. 因此，存在一列初等矩阵 $\boldsymbol{E}_1,\cdots,\boldsymbol{E}_p$ 使得

$$\boldsymbol{A}=\boldsymbol{E}_p\boldsymbol{E}_{p-1}\cdots\boldsymbol{E}_1\cdot\boldsymbol{I}_n=\boldsymbol{E}_p\boldsymbol{E}_{p-1}\cdots\boldsymbol{E}_1$$

简单起见，用 $|\boldsymbol{A}|$ 表示 $\det \boldsymbol{A}$，反复运用重述后的定理 3，我们有：

$$\begin{aligned}|\boldsymbol{AB}|&=|\boldsymbol{E}_p\cdots\boldsymbol{E}_1\boldsymbol{B}|=|\boldsymbol{E}_p||\boldsymbol{E}_{p-1}\cdots\boldsymbol{E}_1\boldsymbol{B}|=\cdots\\&=|\boldsymbol{E}_p|\cdots|\boldsymbol{E}_1||\boldsymbol{B}|=\cdots=|\boldsymbol{E}_p\cdots\boldsymbol{E}_1||\boldsymbol{B}|\\&=|\boldsymbol{A}||\boldsymbol{B}|\end{aligned}$$ ■

基础练习

1. 用最简便的方法计算 $\begin{vmatrix}1&-3&1&-2\\2&-5&-1&-2\\0&-4&5&1\\-3&10&-6&8\end{vmatrix}$.

2. 用一个行列式判断 $\boldsymbol{v}_1$，$\boldsymbol{v}_2$，$\boldsymbol{v}_3$ 是否线性无关，其中，

$$\boldsymbol{v}_1=\begin{bmatrix}5\\-7\\9\end{bmatrix},\qquad \boldsymbol{v}_2=\begin{bmatrix}-3\\3\\-5\end{bmatrix},\qquad \boldsymbol{v}_3=\begin{bmatrix}2\\-7\\5\end{bmatrix}$$

[198]

习题 3.2

习题 1～4 中的每个等式描述了行列式的一个性质. 说出这些性质.

1. $\begin{vmatrix}0&5&-2\\1&-3&6\\4&-1&8\end{vmatrix}=-\begin{vmatrix}1&-3&6\\0&5&-2\\4&-1&8\end{vmatrix}$

2. $\begin{vmatrix}2&-6&4\\3&5&-2\\1&6&3\end{vmatrix}=2\begin{vmatrix}1&-3&2\\3&5&-2\\1&6&3\end{vmatrix}$

3. $\begin{vmatrix}1&3&-4\\2&0&-3\\5&-4&7\end{vmatrix}=\begin{vmatrix}1&3&-4\\0&-6&5\\5&-4&7\end{vmatrix}$

4. $\begin{vmatrix}1&2&3\\0&5&-4\\3&7&4\end{vmatrix}=\begin{vmatrix}1&2&3\\0&5&-4\\0&1&-5\end{vmatrix}$

通过将矩阵行化简为阶梯形式，计算习题5～10中的行列式.

5. $\begin{vmatrix} 1 & 5 & -6 \\ -1 & -4 & 4 \\ -2 & -7 & 9 \end{vmatrix}$ 6. $\begin{vmatrix} 1 & 5 & -3 \\ 3 & -3 & 3 \\ 2 & 13 & -7 \end{vmatrix}$ 7. $\begin{vmatrix} 1 & 3 & 0 & 2 \\ -2 & -5 & 7 & 4 \\ 3 & 5 & 2 & 1 \\ 1 & -1 & 2 & -3 \end{vmatrix}$

8. $\begin{vmatrix} 1 & 3 & 3 & -4 \\ 0 & 1 & 2 & -5 \\ 2 & 5 & 4 & -3 \\ -3 & -7 & -5 & 2 \end{vmatrix}$ 9. $\begin{vmatrix} 1 & -1 & -3 & 0 \\ 0 & 1 & 5 & 4 \\ -1 & 2 & 8 & 5 \\ 3 & -1 & -2 & 3 \end{vmatrix}$ 10. $\begin{vmatrix} 1 & 3 & -1 & 0 & -2 \\ 0 & 2 & -4 & -1 & -6 \\ -2 & -6 & 2 & 3 & 9 \\ 3 & 7 & -3 & 8 & -7 \\ 3 & 5 & 5 & 2 & 7 \end{vmatrix}$

结合使用行化简和余子式展开的方法计算习题11～14中的行列式.

11. $\begin{vmatrix} 2 & 5 & -3 & -1 \\ 3 & 0 & 1 & -3 \\ -6 & 0 & -4 & 9 \\ 4 & 10 & -4 & -1 \end{vmatrix}$ 12. $\begin{vmatrix} -1 & 2 & 3 & 0 \\ 3 & 4 & 3 & 0 \\ 5 & 4 & 6 & 6 \\ 4 & 2 & 4 & 3 \end{vmatrix}$ 13. $\begin{vmatrix} 2 & 5 & 4 & 1 \\ 4 & 7 & 6 & 2 \\ 6 & -2 & -4 & 0 \\ -6 & 7 & 7 & 0 \end{vmatrix}$ 14. $\begin{vmatrix} -3 & -2 & 1 & -4 \\ 1 & 3 & 0 & -3 \\ -3 & 4 & -2 & 8 \\ 3 & -4 & 0 & 4 \end{vmatrix}$

计算习题15～20中的行列式，其中$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}=7$.

15. $\begin{vmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{vmatrix}$ 16. $\begin{vmatrix} a & b & c \\ 3d & 3e & 3f \\ g & h & i \end{vmatrix}$ 17. $\begin{vmatrix} a & b & c \\ g & h & i \\ d & e & f \end{vmatrix}$

18. $\begin{vmatrix} g & h & i \\ a & b & c \\ d & e & f \end{vmatrix}$ 19. $\begin{vmatrix} a & b & c \\ 2d+a & 2e+b & 2f+c \\ g & h & i \end{vmatrix}$ 20. $\begin{vmatrix} a+d & b+e & c+f \\ d & e & f \\ g & h & i \end{vmatrix}$

利用行列式判断习题21～23中的矩阵是否可逆.

21. $\begin{bmatrix} 2 & 3 & 0 \\ 1 & 3 & 4 \\ 1 & 2 & 1 \end{bmatrix}$ 22. $\begin{bmatrix} 5 & 0 & -1 \\ 1 & -3 & -2 \\ 0 & 5 & 3 \end{bmatrix}$ 23. $\begin{bmatrix} 2 & 0 & 0 & 8 \\ 1 & -7 & -5 & 0 \\ 3 & 8 & 6 & 0 \\ 0 & 7 & 5 & 4 \end{bmatrix}$

利用行列式判断习题24～26中的向量集是否线性无关.

24. $\begin{bmatrix} 4 \\ 6 \\ -7 \end{bmatrix}, \begin{bmatrix} -7 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} -3 \\ -5 \\ 6 \end{bmatrix}$ 25. $\begin{bmatrix} 7 \\ -4 \\ -6 \end{bmatrix}, \begin{bmatrix} -8 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 7 \\ 0 \\ -5 \end{bmatrix}$ 26. $\begin{bmatrix} 3 \\ 5 \\ -6 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ -6 \\ 0 \\ 7 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ -3 \end{bmatrix}$

199 在习题27和28中，$\boldsymbol{A}$和$\boldsymbol{B}$都是$n\times n$矩阵. 判断下列命题的真假，并说明理由.

27. a. 行替换变换不改变矩阵的行列式.
    b. $\boldsymbol{A}$的行列式等于$\boldsymbol{A}$的任意阶梯形式$\boldsymbol{U}$的主元之积乘以$(-1)^r$，其中r是将$\boldsymbol{A}$行化简为$\boldsymbol{U}$的过程中行交换的次数.
    c. 如果$\boldsymbol{A}$的列向量线性相关，则$\det \boldsymbol{A}=0$.
    d. $\det(\boldsymbol{A}+\boldsymbol{B})=\det \boldsymbol{A}+\det \boldsymbol{B}$.
28. a. 连做两次行交换，行列式不变.
    b. $\boldsymbol{A}$的行列式等于$\boldsymbol{A}$中对角线元素之积.

c. 如果 $\det A=0$，则 A 包含相同的两行或两列，或者包含一个零行或零列.

d. $\det A^{T}=(-1)\det A$.

29. 计算 $\det B^5$，其中 $B=\begin{bmatrix}1&0&1\\1&1&2\\1&2&1\end{bmatrix}$.

30. 运用定理3(而非定理4)证明：若方阵A包含相同的两行，则 $\det A=0$. 对于 A 的列也有同样的结论，为什么？

31. 证明：如果 A 可逆，则 $\det A^{-1}=\dfrac{1}{\det A}$.

32. 如果 A 是 $n\times n$ 矩阵，给出 $\det(rA)$ 的计算公式.

33. 设 A 和 B 是方阵，证明：$AB\neq BA$ 但 $\det AB=\det BA$.

34. 设 A 和 P 是方阵且 P 可逆，证明 $\det(PAP^{-1})=\det A$.

35. 设 U 是方阵且 $U^{T}U=I$，证明 $\det U=\pm1$.

36. 设 A 是方阵且 $\det A^4=0$，解释 A 为什么不可逆.

验证习题37和38中的矩阵满足 $\det(AB)=(\det A)(\det B)$(不使用定理6).

37. $A=\begin{bmatrix}3&0\\6&1\end{bmatrix}$, $B=\begin{bmatrix}2&0\\5&4\end{bmatrix}$

38. $A=\begin{bmatrix}3&6\\-1&-2\end{bmatrix}$, $B=\begin{bmatrix}4&2\\-1&-1\end{bmatrix}$

39. 设 A 和 B 是 3×3 方阵且 $\det A=4$，$\det B=-3$，利用行列式的性质(包括课文中以及前面习题中提到的性质)计算：

a. $\det AB$　　b. $\det 5A$　　c. $\det B^{T}$　　d. $\det A^{-1}$　　e. $\det A^3$

40. 设 A 和 B 是 4×4 方阵且 $\det A=-1$，$\det B=2$，计算：

a. $\det AB$　　b. $\det B^5$　　c. $\det 2A$　　d. $\det A^{T}A$　　e. $\det B^{-1}AB$

41. 验证 $\det A=\det B+\det C$，其中，

$$A=\begin{bmatrix}a+e&b+f\\c&d\end{bmatrix},\ B=\begin{bmatrix}a&b\\c&d\end{bmatrix},\ C=\begin{bmatrix}e&f\\c&d\end{bmatrix}$$

42. 设 $A=\begin{bmatrix}1&0\\0&1\end{bmatrix}$, $B=\begin{bmatrix}a&b\\c&d\end{bmatrix}$，证明：$\det(A+B)=\det A+\det B$ 当且仅当 $a+d=0$.

43. 验证 $\det A=\det B+\det C$，其中，

$$A=\begin{bmatrix}a_{11}&a_{12}&u_1+v_1\\a_{21}&a_{22}&u_2+v_2\\a_{31}&a_{32}&u_3+v_3\end{bmatrix},\ B=\begin{bmatrix}a_{11}&a_{12}&u_1\\a_{21}&a_{22}&u_2\\a_{31}&a_{32}&u_3\end{bmatrix},\ C=\begin{bmatrix}a_{11}&a_{12}&v_1\\a_{21}&a_{22}&v_2\\a_{31}&a_{32}&v_3\end{bmatrix}$$

注意，$A\neq B+C$.

44. 右乘以初等矩阵 E 在 A 中列上的作用效果与左乘以 E 在 A 中行上的作用效果相同. 利用定理5和定理3以及 E^{T} 也是初等矩阵的事实，证明

$$\det AE=(\det E)(\det A)$$

注意，不要使用定理6.

45. [M] 生成若干 4×5 随机矩阵和 5×6 随机矩阵，对每个矩阵 A，计算 $\det A^{T}A$ 与 $\det AA^{T}$ 的值. 当 A 的列数大于行数时，对于 $A^{T}A$ 和 AA^{T}，你能得出什么结论？

46. [M] 如果 $\det A$ 接近零，矩阵 A 是否也接近奇异矩阵？试用2.3节习题9中近似奇异矩阵的 4×4 方阵 A 验证假设. 计算 A, $10A$ 和 $0.1A$ 的行列式以及条件数，对照计算结果. 设

A 是4×4单位矩阵，重复上述过程．对全部计算结果进行讨论．

基础练习答案

1. 首先，进行行变换使得第一列中多出现零元素，然后会得到某行的元素均为零．

$$\begin{vmatrix} 1 & -3 & 1 & -2 \\ 2 & -5 & -1 & -2 \\ 0 & -4 & 5 & 1 \\ -3 & 10 & -6 & 8 \end{vmatrix} = \begin{vmatrix} 1 & -3 & 1 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & -4 & 5 & 1 \\ 0 & 1 & -3 & 2 \end{vmatrix} = \begin{vmatrix} 1 & -3 & 1 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & -4 & 5 & 1 \\ 0 & 0 & 0 & 0 \end{vmatrix} = 0$$

2. $\det[\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \boldsymbol{v}_3] = \begin{vmatrix} 5 & -3 & 2 \\ -7 & 3 & -7 \\ 9 & -5 & 5 \end{vmatrix} = \begin{vmatrix} 5 & -3 & 2 \\ -2 & 0 & -5 \\ 9 & -5 & 5 \end{vmatrix}$（将第一行元素加到第二行元素的对应位置）

$= -(-3)\begin{vmatrix} -2 & -5 \\ 9 & 5 \end{vmatrix} - (-5)\begin{vmatrix} 5 & 2 \\ -2 & -5 \end{vmatrix}$（按第二列余子式展开）

$= 3 \times 35 + 5 \times (-21) = 0$

由定理4可知，矩阵$[\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \boldsymbol{v}_3]$不可逆．由可逆矩阵定理可知，这些列向量线性相关．

3.3 克莱姆法则、体积与线性变换

本节将运用前几节的理论得到若干重要的理论公式，并给出行列式的一种几何解释．

3.3.1 克莱姆法则

克莱姆法则在许多理论计算中都有应用．例如，可以用它来研究 $\boldsymbol{b}$ 中元素发生变化时，方程 $A\boldsymbol{x}=\boldsymbol{b}$ 解的变化．然而，该公式仅仅适于手算2×2或3×3的小矩阵．

对于任意 $n\times n$ 矩阵 A 以及任意 $\boldsymbol{b}\in\mathbf{R}^n$，设 $A_i(\boldsymbol{b})$ 是以向量 $\boldsymbol{b}$ 代替 A 中第 i 列所得的矩阵，即，

$$A_i(\boldsymbol{b}) = [\boldsymbol{a}_1 \quad \cdots \quad \underset{\substack{\uparrow \\ \text{第 } i \text{ 列}}}{\boldsymbol{b}} \quad \cdots \quad \boldsymbol{a}_n]$$

【定理7】 克莱姆法则

设 A 是一个可逆的 $n\times n$ 矩阵．对任意 $\boldsymbol{b}\in\mathbf{R}^n$，方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 有唯一解：

$$x_i = \frac{\det A_i(\boldsymbol{b})}{\det A}, \qquad i=1,2,\cdots,n \tag{1}$$

证明： 以 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 记 A 的列，以 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n$ 记 $n\times n$ 单位矩阵 I 的列．如果 $A\boldsymbol{x}=\boldsymbol{b}$，则由矩阵乘法的定义可知：

$$\begin{aligned} A\cdot I_i(\boldsymbol{x}) &= A[\boldsymbol{e}_1 \quad \cdots \quad \boldsymbol{x} \quad \cdots \quad \boldsymbol{e}_n] = [A\boldsymbol{e}_1 \quad \cdots \quad A\boldsymbol{x} \quad \cdots \quad A\boldsymbol{e}_n] \\ &= [\boldsymbol{a}_1 \quad \cdots \quad \boldsymbol{b} \quad \cdots \quad \boldsymbol{a}_n] = A_i(\boldsymbol{b}) \end{aligned}$$

由行列式的乘法性质可知：

$$(\det A)(\det I_i(\boldsymbol{x})) = \det A_i(\boldsymbol{b})$$

其中，等式左端的第二个行列式就是 x_i（按第 i 行展开余子式即可得到）．因此，$(\det A)\cdot x_i = \det A_i(\boldsymbol{b})$．由于 A 可逆且 $\det A\neq 0$，(1)式得证． ■

【例题1】 运用克莱姆法则求解方程组：

$$3x_1 - 2x_2 = 6$$
$$-5x_1 + 4x_2 = 8$$

解：将此方程组看作 $\boldsymbol{Ax}=\boldsymbol{b}$. 使用前面介绍的记号，我们有

$$\boldsymbol{A}=\begin{bmatrix}3 & -2\\ -5 & 4\end{bmatrix},\qquad \boldsymbol{A}_1(\boldsymbol{b})=\begin{bmatrix}6 & -2\\ 8 & 4\end{bmatrix},\qquad \boldsymbol{A}_2(\boldsymbol{b})=\begin{bmatrix}3 & 6\\ -5 & 8\end{bmatrix}$$

由于 $\det \boldsymbol{A}=2$，根据克莱姆法则，此方程组有唯一解：

$$x_1=\frac{\det \boldsymbol{A}_1(\boldsymbol{b})}{\det \boldsymbol{A}}=\frac{24+16}{2}=20$$
$$x_2=\frac{\det \boldsymbol{A}_2(\boldsymbol{b})}{\det \boldsymbol{A}}=\frac{24+30}{2}=27$$ ■

3.3.2 工程应用

许多重要的工程问题，特别是电气工程和控制理论中的问题，都可以利用拉普拉斯变换进行分析. 这种方法将一个线性微分方程组转换为系数含参数的线性代数方程组. 下面的例题就给出了这样的一个代数方程组.

【例题 2】 考虑下列方程组，其中 s 是待定的参数. 运用克莱姆法则确定当方程组有唯一解时，s 的值是多少，并求出方程组的解.

$$3sx_1 - 2x_2 = 4$$
$$-6x_1 + sx_2 = 1$$

解：将此方程组看作 $\boldsymbol{Ax}=\boldsymbol{b}$ 的形式. 则

$$\boldsymbol{A}=\begin{bmatrix}3s & -2\\ -6 & s\end{bmatrix},\qquad \boldsymbol{A}_1(\boldsymbol{b})=\begin{bmatrix}4 & -2\\ 1 & s\end{bmatrix},\qquad \boldsymbol{A}_2(\boldsymbol{b})=\begin{bmatrix}3s & 4\\ -6 & 1\end{bmatrix}$$

202

由于

$$\det \boldsymbol{A}=3s^2-12=3(s+2)(s-2)$$

根据克莱姆法则，仅当 $s\neq \pm 2$ 时，该方程组有唯一解. 对这样的 s，解为 (x_1,x_2)，其中

$$x_1=\frac{\det \boldsymbol{A}_1(\boldsymbol{b})}{\det \boldsymbol{A}}=\frac{4s+2}{3(s+2)(s-2)}$$
$$x_2=\frac{\det \boldsymbol{A}_2(\boldsymbol{b})}{\det \boldsymbol{A}}=\frac{3s+24}{3(s+2)(s-2)}=\frac{s+8}{(s+2)(s-2)}$$ ■

3.3.3 A^{-1}的计算公式

运用克莱姆法则很容易得到求 $n\times n$ 矩阵 $\boldsymbol{A}$ 的逆矩阵的一般公式. $\boldsymbol{A}^{-1}$第 j 列是满足下述公式的向量 $\boldsymbol{x}$：

$$\boldsymbol{Ax}=\boldsymbol{e}_j$$

其中 $\boldsymbol{e}_j$是单位矩阵的第 j 列，$\boldsymbol{x}$ 的第 i 个元素是 $\boldsymbol{A}^{-1}$的 (i,j) – 元. 由克莱姆法则，

$$\{\boldsymbol{A}^{-1}\text{的}(i,j)-\text{元}\}=x_i=\frac{\det \boldsymbol{A}_i\ (\boldsymbol{e}_j)}{\det \boldsymbol{A}} \tag{2}$$

回忆记号 $\boldsymbol{A}_{ji}$表示划去 $\boldsymbol{A}$ 中第 j 行、第 i 列所得的矩阵，观察 $\boldsymbol{A}_i(\boldsymbol{e}_j)$按第 i 列的余子式展开可知

$$\det A_i(e_j) = (-1)^{i+j}\det A_{ji} = C_{ji} \tag{3}$$

其中 C_{ji}是 A 的余子式. 由公式(2)可知，A^{-1}的(i, j)－元是 $\det A$ 除 C_{ji}的商. [注意，C_{ji}的下标是(i,j)的逆序]. 因此

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} \tag{4}$$

(4)式右端由余子式构成的矩阵称作 A 的转置**伴随矩阵**(adjugate)或**经典伴随矩阵**(classical adjoint)(术语伴随矩阵在线性变换的高等教材中有其他含义). 下面的定理简单重述了公式(4).

【定理8】　求逆公式

设 A 是一个可逆的 $n\times n$ 矩阵，则

$$A^{-1} = \frac{1}{\det A}\text{adj}\, A$$

203

【例题3】　计算矩阵 $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix}$的逆.

解: 9个余子式分别是:

$$C_{11} = +\begin{vmatrix} -1 & 1 \\ 4 & -2 \end{vmatrix} = -2,\ C_{12} = -\begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix} = 3,\ C_{13} = +\begin{vmatrix} 1 & -1 \\ 1 & 4 \end{vmatrix} = 5$$

$$C_{21} = -\begin{vmatrix} 1 & 3 \\ 4 & -2 \end{vmatrix} = 14,\ C_{22} = +\begin{vmatrix} 2 & 3 \\ 1 & -2 \end{vmatrix} = -7,\ C_{23} = -\begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = -7$$

$$C_{31} = +\begin{vmatrix} 1 & 3 \\ -1 & 1 \end{vmatrix} = 4,\ C_{32} = -\begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = 1,\ C_{33} = +\begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} = -3$$

伴随矩阵是这些余子式所构成矩阵的转置(例如，C_{12}位于伴随矩阵$(2,1)$位置). 因此有:

$$\text{adj}\, A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}$$

我们可以直接计算 $\det A$，但以下计算可以帮助我们检验上面的计算结果，同时求出 $\det A$:

$$(\text{adj}\, A)\cdot A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix} = 14I$$

由于$(\text{adj}\, A)A = 14I$，由定理8得 $\det A = 14$，且

$$A^{-1}=\frac{1}{14}\begin{bmatrix}-2 & 14 & 4\\ 3 & -7 & 1\\ 5 & -7 & -3\end{bmatrix}=\begin{bmatrix}-1/7 & 1 & 2/7\\ 3/14 & -1/2 & 1/14\\ 5/14 & -1/2 & -3/14\end{bmatrix}$$ ■

注记

定理 8 主要应用于理论计算中．它所给出的 A^{-1} 公式使得我们不计算逆矩阵也能推导出它的性质．在需要具体计算 A^{-1} 时，除几种特殊情况以外，2.2 节给出的算法是较好的一种方法．

克莱姆法则也是一个理论计算工具．可用于研究方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的解对 $\boldsymbol{b}$ 或者 A 中元素变动（多半由获取 $\boldsymbol{b}$ 或者 A 中元素的实验误差引起）的敏感程度．设 A 是 3×3 的复数矩阵，由于 $[A \quad \boldsymbol{b}]$ 的行化简涉及复数运算，可能较复杂，而行列式则相对容易计算，因此我们有时会使用克莱姆法则做手算．对于 $n\times n$（实数或复数）矩阵，克莱姆法则几乎无法使用，因为一个行列式的计算量与利用行化简求解 $A\boldsymbol{x}=\boldsymbol{b}$ 一样多．

3.3.4 面积或体积的行列式解释

在接下来的应用里，我们将验证本章实例介绍中提到的行列式的几何解释．尽管直到第 6 章才会专门讨论 $\mathbf{R}^n$ 中的长度和距离，此处假定大家已经理解并掌握了 $\mathbf{R}^2$ 和 $\mathbf{R}^3$ 中的欧几里德长度、面积和体积． 204

【定理 9】 如果 A 是 2×2 矩阵，则由 A 的列向量确定的平行四边形的面积等于 $|\det A|$．如果 A 是 3×3 矩阵，则由 A 的列向量确定的平行六面体的体积等于 $|\det A|$．

证明： 对于 2×2 对角矩阵，定理显然成立：

$$\left|\det\begin{bmatrix}a & 0\\ 0 & d\end{bmatrix}\right| = |ad| = \{\text{矩形的面积}\}$$

如图 3-4 所示．我们只需证明，任意 2×2 矩阵 $A=[\boldsymbol{a}_1 \quad \boldsymbol{a}_2]$ 可以以某种方式变换成对角矩阵，既不改变其对应平行四边形的面积，也不改变 $|\det A|$．由 3.2 节可知，交换两列或者将一列的倍数加到另一列，行列式的绝对值不变．容易看出，这样的变换可将任意 A 转换为对角矩阵．而交换两列不会改变平行四边形．所以，只需验证 $\mathbf{R}^2$ 或 $\mathbf{R}^3$ 中的向量满足下列几何性质：

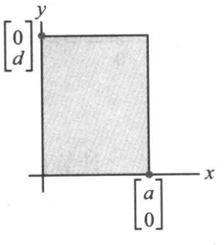


图 3-4　面积 $=|ad|$

设 $\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2$ 是两个非零向量．对任意数量 c，由 $\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2$ 确定的平行四边形的面积等于由 $\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2+c\boldsymbol{a}_1$ 确定的平行四边形的面积．

为证明这个命题，假定 $\boldsymbol{a}_2$ 不是 $\boldsymbol{a}_1$ 的倍数，否则两个平行四边形都将退化且面积都为零．若 L 是过 $\mathbf{0}$ 和 $\boldsymbol{a}_1$ 的直线，则 $\boldsymbol{a}_2+L$ 是经过 $\boldsymbol{a}_2$ 且平行于 L 的直线，且 $\boldsymbol{a}_2+c\boldsymbol{a}_1$ 在这条直线上，如图 3-5 所示，点 $\boldsymbol{a}_2$ 和点 $\boldsymbol{a}_2+c\boldsymbol{a}_1$ 到直线 L 的垂直距离相等．由于两平行四边形有公共的底，即 $\mathbf{0}$ 到 $\boldsymbol{a}_1$ 的线段，因此图 3-5 中两个平行四边形的面积相

同. 这就证明了 $\mathbf{R}^2$ 的情形.

$\mathbf{R}^3$ 的证明是类似的. 对于任意 3×3 对角阵，定理显然成立. 见图 3-6. 任意 3×3 矩阵 $\boldsymbol{A}$ 都可以通过列变换变成对角矩阵，并且保证 $\det \boldsymbol{A}$ 不变(考虑对 $\boldsymbol{A}^{\mathrm{T}}$ 做行变换). 因此，只需证明这些变换不会改变 $\boldsymbol{A}$ 的列所确定的平行六面体的体积.

205

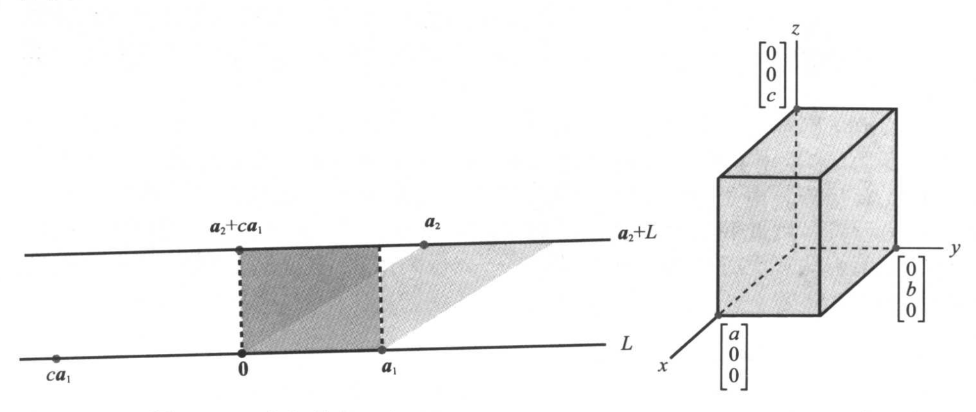


图 3-5　面积相等的两个平行四边形　　图 3-6　体积 = $|abc|$

图 3-7 所示的平行六面体是一个含两个倾斜面的阴影体. 其体积等于 $\mathrm{Span}\{\boldsymbol{a}_1,\boldsymbol{a}_3\}$ 平面上的底面积乘以 $\boldsymbol{a}_2$ 在 $\mathrm{Span}\{\boldsymbol{a}_1,\boldsymbol{a}_3\}$ 上的高. 任意 $\boldsymbol{a}_2+c\boldsymbol{a}_1$ 形式的向量高度相同，因为它们同属于平面 $\boldsymbol{a}_2+\mathrm{Span}\{\boldsymbol{a}_1,\boldsymbol{a}_3\}$. 因此当 $[\boldsymbol{a}_1\quad \boldsymbol{a}_2\quad \boldsymbol{a}_3]$ 变成 $[\boldsymbol{a}_1\quad \boldsymbol{a}_2+c\boldsymbol{a}_1\quad \boldsymbol{a}_3]$ 时，平行六面体的体积保持不变. 由于交换两列对体积没有任何影响，定理得证. ■

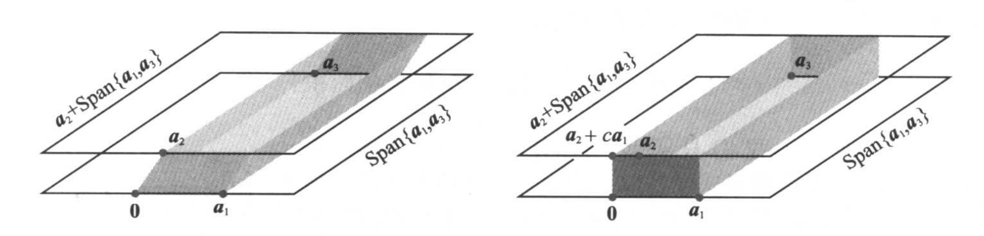


图 3-7　面积相等的两个平行六面体

【例题 4】　计算由点 $(-2,-2)$，$(0,3)$，$(4,-1)$ 与 $(6,4)$ 所确定的平行四边形的面积，如图 3-8(a) 所示.

解： 首先平移该平行四边形，使其有一顶点为原点. 例如，给四个顶点的坐标都减去 $(-2,-2)$. 这个新的平行四边形与原平行四边形面积相等，其顶点的坐标分别为 $(0,0)$，$(2,5)$，$(6,1)$ 与 $(8,6)$，观察图 3-8(b)，该平行四边形由下列矩阵确定：

$$\boldsymbol{A}=\begin{bmatrix}2&6\\5&1\end{bmatrix}$$

由于 $|\det \boldsymbol{A}|=|-28|$，平行四边形的面积应为 28. ■

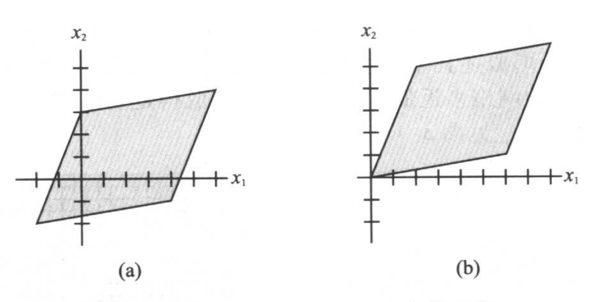


图 3-8 平移平行四边形不改变其面积

206

3.3.5 线性变换

行列式可用于描述平面以及 $\mathbf{R}^3$ 中线性变换的一个重要的几何性质. 若 T 是一个线性变换，S 是 T 的定义域内的一个集合，令 $T(S)$ 代表 S 中点的像所构成的集合. 我们关心的是 $T(S)$ 的面积（或体积）与原来 S 的面积（或体积）之间的联系. 为简便起见，若 S 是一个平行四边形围成的区域，我们就说 S 是一个平行四边形.

【定理 10】 设 T: $\mathbf{R}^2 \to \mathbf{R}^2$ 是由 2×2 矩阵 A 确定的一个线性变换，若 S 是 $\mathbf{R}^2$ 中的一个平行四边形，则

$$\{T(S)\text{的面积}\} = |\det A| \cdot \{S\text{的面积}\} \tag{5}$$

若 T 是由 3×3 矩阵 A 确定的线性变换且 S 是 $\mathbf{R}^3$ 中的一个平行六面体，则

$$\{T(S)\text{的体积}\} = |\det A| \cdot \{S\text{的体积}\} \tag{6}$$

证明： 考虑 2×2 矩阵 $A = [\boldsymbol{a}_1 \quad \boldsymbol{a}_2]$. $\mathbf{R}^2$ 中一个顶点在原点且由向量 $\boldsymbol{b}_1$ 和 $\boldsymbol{b}_2$ 确定的平行四边形为：

$$S = \{s_1\boldsymbol{b}_1 + s_2\boldsymbol{b}_2 : 0 \leqslant s_1 \leqslant 1, 0 \leqslant s_2 \leqslant 1\}$$

S 在 T 下的像由如下形式的点构成：

$$T(s_1\boldsymbol{b}_1 + s_2\boldsymbol{b}_2) = s_1T(\boldsymbol{b}_1) + s_2T(\boldsymbol{b}_2) = s_1A\boldsymbol{b}_1 + s_2A\boldsymbol{b}_2$$

其中 $0 \leqslant s_1 \leqslant 1$，$0 \leqslant s_2 \leqslant 1$. 因此，$T(S)$ 是由矩阵 $[A\boldsymbol{b}_1 \quad A\boldsymbol{b}_2]$ 的列所确定的平行四边形. 该矩阵可写成 AB，其中 $B = [\boldsymbol{b}_1 \quad \boldsymbol{b}_2]$. 由定理 9 以及行列式的乘法性质可知，

$$\begin{aligned}\{T(S)\text{的面积}\} &= |\det AB| = |\det A| \cdot |\det B| \\ &= |\det A| \cdot \{S\text{的面积}\}\end{aligned} \tag{7}$$

任意的平行四边形都有 $\boldsymbol{p}+S$ 的形式，其中 $\boldsymbol{p}$ 是一个向量，S 是上文中一个顶点在原点的平行四边形. 容易看到，线性变换 T 将 $\boldsymbol{p}+S$ 映到 $T(\boldsymbol{p})+T(S)$（见习题 26）. 由于平移不改变集合的面积，因此，

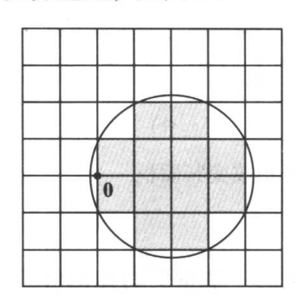
$$\begin{aligned}\{T(\boldsymbol{p}+S)\text{的面积}\} &= \{T(\boldsymbol{p})+T(S)\text{的面积}\} \\ &= \{T(S)\text{的面积}\} && \text{平移} \\ &= |\det A| \cdot \{S\text{的面积}\} && \text{根据(7)} \\ &= |\det A| \cdot \{\boldsymbol{p}+S\text{的面积}\} && \text{平移}\end{aligned}$$

这就证明了(5)对于 $\mathbf{R}^2$ 中的全体平行四边形都成立. 类似地，我们可以考察 3×3 的情形并且证明(6). ■

207

要把定理 10 推广到 $\mathbf{R}^2$ 或者 $\mathbf{R}^3$ 中并非由直线或者平面围成的某个区域，我们将

面临这样一个问题：如何定义并计算它的面积或体积. 这属于微积分的研究范围，这里我们只对 $\boldsymbol{R}^2$ 的情形概述其基本思想. 若 R 是面积有限的平面区域，我们可以用 R 中小正方形所构成的网格来近似 R. 当正方形选取得足够小时，R 的面积可以由它们面积的总和充分逼近，见图 3-9.



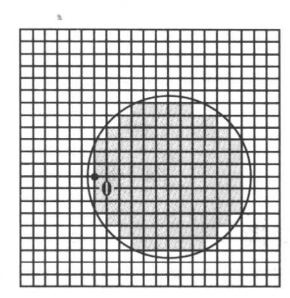


图 3-9 用小正方形的并逼近平面区域. 网格越小逼近效果越好

若 T 是 2×2 矩阵 $\boldsymbol{A}$ 所确定的线性变换，平面区域 R 在 T 下的像可以用 R 中小正方形的像来逼近. 定理 10 的证明说明，每个小正方形的像都是一个平行四边形，其面积等于正方形的面积乘以 $|\det \boldsymbol{A}|$. 设 R' 是 R 中正方形的并，则 $T(R')$ 的面积等于 R' 的面积乘以 $|\det \boldsymbol{A}|$，见图 3-10. $T(R')$ 的面积就是 $T(R)$ 面积的一个逼近.

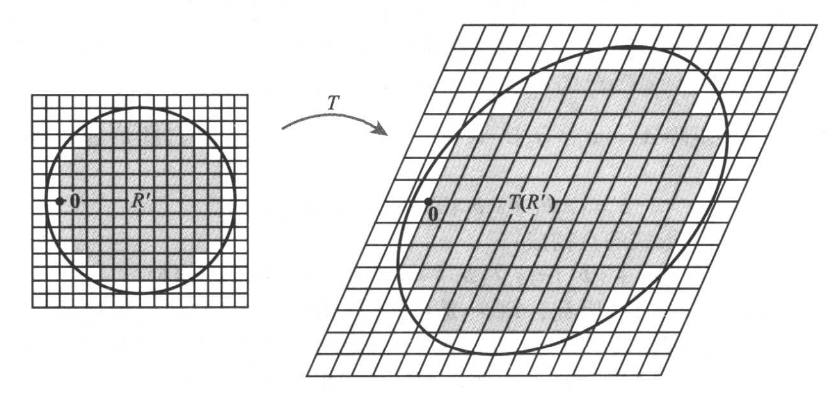


图 3-10 用平行四边形的并逼近 $T(R)$

208 在上述讨论中引入极限，就能证明定理 10 的下述推广：

当 S 是 $\mathbf{R}^2$ 中面积有限的区域或者是 $\mathbf{R}^3$ 中体积有限的区域时，定理 10 的结论也成立.

【例题 5】 设 a 和 b 都是正数. 计算区域 E 的面积，其边界是下列方程确定的椭圆：

$$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}=1$$

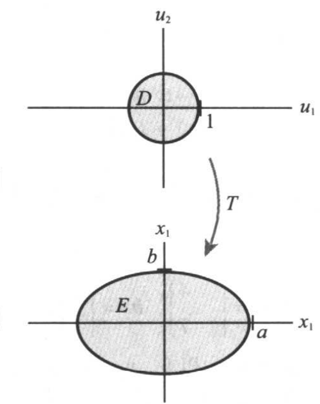
解： 可以断定 E 是单位圆盘在矩阵 $\boldsymbol{A}=\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ 所确定的线性变换 T 之下的像. 因为

如果 $\boldsymbol{u}=\begin{bmatrix}u_1\\u_2\end{bmatrix}$，$\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\end{bmatrix}$，且$\boldsymbol{x}=A\boldsymbol{u}$，则有

$$u_1=\frac{x_1}{a},u_2=\frac{x_2}{b}$$

于是，$\boldsymbol{u}$ 在单位圆盘内且 $u_1^2+u_2^2\leqslant 1$ 当且仅当 $\boldsymbol{x}$ 在 E 内且 $(x_1/a)^2+(x_2/b)^2\leqslant 1$. 由定理 10 的推广可知，

$$\begin{aligned}\{椭圆面积\}&=\{T(D)的面积\}\\&=|\det A|\cdot\{D\text{ 的面积}\}\\&=ab\cdot\pi(1)^2=\pi ab\end{aligned}$$ ■



基础练习

设 S 是由向量 $\boldsymbol{b}_1=\begin{bmatrix}1\\3\end{bmatrix}$和 $\boldsymbol{b}_2=\begin{bmatrix}5\\1\end{bmatrix}$确定的平行四边形，且 $A=\begin{bmatrix}1&-0.1\\0&2\end{bmatrix}$. 计算 S 在映射 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 下像的面积.

习题 3.3

运用克莱姆法则计算习题 1 ~ 6 中各个方程组的解.

1. $\begin{aligned}5x_1+7x_2&=3\\2x_1+4x_2&=1\end{aligned}$

2. $\begin{aligned}4x_1+x_2&=6\\5x_1+2x_2&=7\end{aligned}$

3. $\begin{aligned}3x_1-2x_2&=7\\-5x_1+6x_2&=-5\end{aligned}$

4. $\begin{aligned}-5x_1+3x_2&=9\\3x_1-x_2&=-5\end{aligned}$

5. $\begin{aligned}2x_1+x_2\qquad&=7\\-3x_1+\qquad x_3&=-8\\x_2+2x_3&=-3\end{aligned}$

6. $\begin{aligned}2x_1+x_2+x_3&=4\\-x_1+\qquad 2x_3&=2\\3x_1+x_2+3x_3&=-2\end{aligned}$

209

在习题 7 ~ 10 中，确定参数 s 的值，使方程组有唯一解，并且求出这个解.

7. $\begin{aligned}6sx_1+4x_2&=5\\9x_1+2sx_2&=-2\end{aligned}$

8. $\begin{aligned}3sx_1-5x_2&=3\\9x_1+5sx_2&=2\end{aligned}$

9. $\begin{aligned}sx_1-2sx_2&=-1\\3x_1+6sx_2&=4\end{aligned}$

10. $\begin{aligned}2sx_1+x_2&=1\\3sx_1+6sx_2&=2\end{aligned}$

计算习题 11 ~ 16 给定矩阵的伴随矩阵，并运用定理 8 求解其逆矩阵.

11. $\begin{bmatrix}0&-2&-1\\3&0&0\\-1&1&1\end{bmatrix}$

12. $\begin{bmatrix}1&1&3\\2&-2&1\\0&1&0\end{bmatrix}$

13. $\begin{bmatrix}3&5&4\\1&0&1\\2&1&1\end{bmatrix}$

14. $\begin{bmatrix}3&6&7\\0&2&1\\2&3&4\end{bmatrix}$

15. $\begin{bmatrix}3&0&0\\-1&1&0\\-2&3&2\end{bmatrix}$

16. $\begin{bmatrix}1&2&4\\0&-3&1\\0&0&3\end{bmatrix}$

17. 证明：若 A 是 2×2 矩阵，则定理 8 中 A^{-1}的公式与 2.2 节的定理 4 中公式一样.

18. 假设 A 中所有元素都是整数且 $\det A=1$. 解释为什么 A^{-1}中所有元素也都是整数.

在习题 19 ~ 20 中，求以给定点为定点的平行六面体的面积.

19. (0, 0)，(5,2)，(6,4)，(11,6)

20. (0,0)，(−1,3)，(4,−5)，(3,−2)

21. (−1,0)，(0,5)，(1,−4)，(2,1)

22. (0,−2)，(6,−1)，(−3,1)，(3,2)

23. 计算以原点，(1,0,−2)，(1,2,4)和(7,1,0)为顶点的平行六面体的体积.

24. 计算以原点，$(1,4,0)$，$(-2,-5,2)$和$(-1,2,-1)$为顶点的平行六面体的体积.
25. 运用体积的概念解释为什么 3×3 矩阵 $\boldsymbol{A}$ 的行列式为零当且仅当 $\boldsymbol{A}$ 不可逆. 此处不要使用 3.2 节的定理 4. [提示：考虑矩阵 $\boldsymbol{A}$ 的列.]
26. 设 T 是 $\mathbf{R}^m\to\mathbf{R}^n$ 的线性变换，$\boldsymbol{p}$ 是向量，S 是 $\mathbf{R}^m$ 的一个集合. 证明：$\boldsymbol{p}+S$ 在 T 下的像是 $\mathbf{R}^n$ 中集合 $T(\boldsymbol{p})+T(S)$.
27. 设 S 是由向量 $\boldsymbol{b}_1=\begin{bmatrix}-2\\3\end{bmatrix}$和 $\boldsymbol{b}_2=\begin{bmatrix}-2\\5\end{bmatrix}$确定的平行四边形，且 $\boldsymbol{A}=\begin{bmatrix}6&-2\\-3&2\end{bmatrix}$. 计算 S 在映射$\boldsymbol{x}\mapsto\boldsymbol{A}\boldsymbol{x}$下像的面积.
28. 令 $\boldsymbol{b}_1=\begin{bmatrix}4\\-7\end{bmatrix}$，$\boldsymbol{b}_2=\begin{bmatrix}0\\1\end{bmatrix}$，且 $\boldsymbol{A}=\begin{bmatrix}7&2\\1&1\end{bmatrix}$. 重做习题 27.
29. 给出 $\mathbf{R}^2$ 中以 $\mathbf{0}$，$\boldsymbol{v}_1$，$\boldsymbol{v}_2$为顶点的三角形的面积公式.
30. 设 R 是以(x_1,y_1)，(x_2,y_2)和(x_3,y_3)为顶点的三角形. 证明：

$$\{三角形的面积\}=\frac{1}{2}\left|\det\begin{bmatrix}x_1&y_1&1\\x_2&y_2&1\\x_3&y_3&1\end{bmatrix}\right|$$

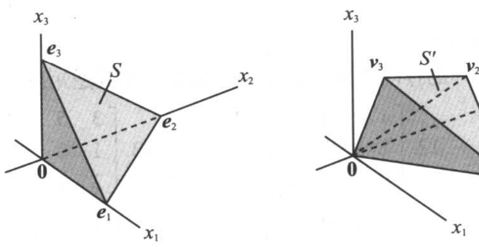
[提示：将各顶点坐标减去某个顶点的坐标，使得 R 平移到原点，然后利用习题 29 的结论.]

31. 设 T：$\mathbf{R}^3\to\mathbf{R}^3$ 是矩阵 $\boldsymbol{A}=\begin{bmatrix}a&0&0\\0&b&0\\0&0&c\end{bmatrix}$确定的线性变换，其中 a，b，c 是正数. 设 S 是单位球体，且边界方程为：$x_1^2+x_2^2+x_3^2=1$.

    a. 证明：$T(S)$的边界是方程$\dfrac{x_1^2}{a^2}+\dfrac{x_2^2}{b^2}+\dfrac{x_3^2}{c^2}=1$ 确定的椭球.

    b. 已知单位球体的体积为 $4\pi/3$，计算(a)中椭球所围成区域的体积.

32. 设 S 是 $\mathbf{R}^3$ 中以 $\mathbf{0}$，$\boldsymbol{e}_1$，$\boldsymbol{e}_2$，$\boldsymbol{e}_3$为顶点的四面体，S'是 $\mathbf{R}^3$ 中以 $\mathbf{0}$，$\boldsymbol{v}_1$，$\boldsymbol{v}_2$，$\boldsymbol{v}_3$为顶点的四面体，如下图所示：



    a. 给出将 S 映到 S'的线性变换 T.

    b. 利用下列事实求四面体 S'的体积公式.

$$\{S 的体积\}=(1/3)\{底面积\}\cdot\{高\}$$

210

33. [M]对一个 4×4 随机矩阵 $\boldsymbol{A}$，验证定理 8 所给的求逆公式. 利用矩阵软件计算 3×3 子矩阵的余子式，构造转置伴随矩阵，令 $\boldsymbol{B}=(\text{adj}\ \boldsymbol{A})/(\det\boldsymbol{A})$. 然后计算 $\boldsymbol{B}-\text{inv}(\boldsymbol{A})$，其中 $\text{inv}(\boldsymbol{A})$是用矩阵软件求得的逆矩阵. 利用浮点运算，尽可能保留最多的小数位. 最后报告你的结论.
34. [M] 用 4×4 随机矩阵 $\boldsymbol{A}$ 和 4×1 随机向量 $\boldsymbol{b}$ 验证克莱姆法则，计算 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 解的每一个元

素，与 $A^{-1}b$ 的元素作比较. 写出你所用的矩阵软件中，运用克莱姆法则求 x 中第二个元素的命令(或键盘输入).

35. [M] 如果你所用 MATLAB 的版本有 flop 命令，可以用它得到计算 30×30 随机矩阵的逆所需的浮点运算次数，将这个值与计算 $(\text{adj}\,A)/(\det A)$ 所需的浮点运算次数作比较.

基础练习答案

S 的面积等于 $\left|\det\begin{bmatrix}1 & 5\\ 3 & 1\end{bmatrix}\right|=14$，且 $\det A=2$. 由定理 10，S 在映射 $x\mapsto Ax$ 下像的面积是：

$$|\det A|\cdot\{S\text{ 的面积}\}=2\times 14=28$$

第 3 章补充题

1. 判断下列命题的真假，并说明理由. 此处假定矩阵皆为方阵.
   a. 若 2×2 矩阵 A 的行列式非零，则 A 中两列成比例.
   b. 若 3×3 矩阵 A 中两行相同，则 $\det A=0$.
   c. 若 A 是 3×3 矩阵， 则 $\det 5A=5\det A$.
   d. 若 A 和 B 是 $n\times n$ 矩阵，且 $\det A=2$，$\det B=3$，则 $\det(A+B)=5$.
   e. 若 A 是 $n\times n$ 矩阵，且 $\det A=2$，则 $\det A^3=6$.
   f. 若交换 A 中两行得到 B，则 $\det B=\det A$.
   g. 若 A 的第三行乘 5 得到 B，则 $\det B=5\det A$.
   h. 若将其余行的一个线性组合加到某一行上得到 B，则 $\det A=\det B$.
   i. $\det A^{\mathrm{T}}=-\det A$.
   j. $\det(-A)=-\det A$.
   k. $\det A^{\mathrm{T}}A\geqslant 0$.
   l. 任意包含 n 个变量、n 个线性方程的方程组都可以用克莱姆法则求解.
   m. 若 u，$v\in\mathbf{R}^2$ 且 $\det[u\ \ v]=10$，则坐标系中以 $\mathbf{0}$，u 和 v 为顶点的三角形的面积是 10.
   n. 若 $A^3=\mathbf{0}$，则 $\det A=0$.
   o. 若 A 可逆，则 $\det A^{-1}=\det A$.
   p. 若 A 可逆，则 $(\det A^{-1})(\det A)=1$.

利用行变换证明习题 2 ~ 4 中的行列式均为 0.

2. $\begin{vmatrix}12 & 13 & 14\\ 15 & 16 & 17\\ 18 & 19 & 20\end{vmatrix}$　3. $\begin{vmatrix}1 & a & b+c\\ 1 & b & a+c\\ 1 & c & a+b\end{vmatrix}$　4. $\begin{vmatrix}a & b & c\\ a+x & b+x & c+x\\ a+y & b+y & c+y\end{vmatrix}$

计算习题 5 ~ 6 的行列式.

5. $\begin{vmatrix}9 & 1 & 9 & 9 & 9\\ 9 & 0 & 9 & 9 & 2\\ 4 & 0 & 0 & 5 & 0\\ 9 & 0 & 3 & 9 & 0\\ 6 & 0 & 0 & 7 & 0\end{vmatrix}$　6. $\begin{bmatrix}4 & 8 & 8 & 8 & 5\\ 0 & 1 & 0 & 0 & 0\\ 6 & 8 & 8 & 8 & 7\\ 0 & 8 & 8 & 3 & 0\\ 0 & 8 & 2 & 0 & 0\end{bmatrix}$

211

7. 证明：$\mathbf{R}^2$ 中过相异两点 $(x_1,\ y_1)$，(x_2,y_2) 的直线方程可以写为：

$$\det\begin{bmatrix}1 & x & y\\ 1 & x_1 & y_1\\ 1 & x_2 & y_2\end{bmatrix}=0$$

8. 求类似于习题7的3×3行列式方程，它给出了过点(x_1,y_1)且斜率为m的直线方程.

习题9和10所讨论的矩阵是下述范德蒙矩阵的行列式：

$$\boldsymbol{T}=\begin{bmatrix}1 & a & a^2\\ 1 & b & b^2\\ 1 & c & c^2\end{bmatrix},\qquad V(t)=\begin{bmatrix}1 & t & t^2 & t^3\\ 1 & x_1 & x_1^2 & x_1^3\\ 1 & x_2 & x_2^2 & x_2^3\\ 1 & x_3 & x_3^2 & x_3^3\end{bmatrix}$$

9. 运用行变换证明：

$$\det \boldsymbol{T}=(b-a)(c-a)(c-b)$$

10. 设$f(t)=\det \boldsymbol{V}$，且x_1，x_2，x_3互不相等. 解释$f(t)$为什么是三次多项式，证明t^3的系数非零，并且求f图像上的三点.
11. 计算以(1,4)，(−1,5)，(3,9)和(5,8)为顶点的平行四边形的面积. 如何判定该四边形就是一个平行四边形？
12. 利用平行四边形面积的概念写出一个关于2×2矩阵$\boldsymbol{A}$的命题，使得命题成立当且仅当$\boldsymbol{A}$可逆.
13. 证明：若矩阵$\boldsymbol{A}$可逆，则adj $\boldsymbol{A}$也可逆，且

$$(\text{adj}\,\boldsymbol{A})^{-1}=\frac{1}{\det \boldsymbol{A}}\boldsymbol{A}$$

[提示：已知矩阵$\boldsymbol{B}$和$\boldsymbol{C}$，怎样证明$\boldsymbol{C}$是$\boldsymbol{B}$的逆?]

14. 设$\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$，$\boldsymbol{D}$与$\boldsymbol{I}$是$n\times n$矩阵. 利用行列式的定义和性质证明以下公式. (c)在特征值(第5章)的应用中很有用.

a. $\det\begin{bmatrix}\boldsymbol{A} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}=\det \boldsymbol{A}$ b. $\det\begin{bmatrix}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{C} & \boldsymbol{D}\end{bmatrix}=\det \boldsymbol{D}$ c. $\det\begin{bmatrix}\boldsymbol{A} & \boldsymbol{0}\\ \boldsymbol{C} & \boldsymbol{D}\end{bmatrix}=(\det \boldsymbol{A})(\det \boldsymbol{D})=\det\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{0} & \boldsymbol{D}\end{bmatrix}$

15. 设$\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$与$\boldsymbol{D}$是$n\times n$矩阵，其中$\boldsymbol{A}$可逆.

a. 求可以产生下列分块LU分解的矩阵$\boldsymbol{X}$和$\boldsymbol{Y}$：

$$\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{bmatrix}=\begin{bmatrix}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{X} & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{0} & \boldsymbol{Y}\end{bmatrix}$$

然后证明：

$$\det\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{bmatrix}=(\det \boldsymbol{A})\cdot\det(\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B})$$

b. 证明：若$\boldsymbol{AC}=\boldsymbol{CA}$，则

$$\det\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{bmatrix}=\det(\boldsymbol{AD}-\boldsymbol{CB})$$

16. 设$\boldsymbol{J}$是元素全为1的$n\times n$矩阵，考虑$\boldsymbol{A}=(a-b)\boldsymbol{I}+b\boldsymbol{J}$，即

$$\boldsymbol{A}=\begin{bmatrix}a & b & b & \cdots & b\\ b & a & b & \cdots & b\\ b & b & a & \cdots & b\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ b & b & b & \cdots & a\end{bmatrix}$$

按照下列步骤验证 $\det \boldsymbol{A}=(a-b)^{n-1}[a+(n-1)b]$：

a. 从第一行中减去第二行，从第二行中减去第三行，依此类推. 解释矩阵的行列式为什么不变.

b. 用(a)得到的矩阵，做下列变换：将第一列加到第二列上，将第二列加到第三列上，依此类推. 解释矩阵的行列式为什么不变.

c. 计算(b)所得矩阵的行列式.

17. 设 $\boldsymbol{A}$ 为习题 16 中的初始矩阵，且

$$\boldsymbol{B}=\begin{bmatrix} a-b & b & b & \cdots & b \\ 0 & a & b & \cdots & b \\ 0 & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & b & b & \cdots & a \end{bmatrix} \quad \boldsymbol{C}=\begin{bmatrix} b & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{bmatrix}$$

注意：此处 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 几乎一样. 但 $\boldsymbol{A}$ 的第一列元素等于 $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 第一列元素之和. 由 3.2 节中提到的行列式函数的线性性质可知，$\det \boldsymbol{A}=\det \boldsymbol{B}+\det \boldsymbol{C}$. 利用此公式，对矩阵的维度作归纳，证明习题 16 的公式. 212

18. [M]利用习题 16 的结果计算下列矩阵的行列式，然后用矩阵软件加以验证.

$$\begin{bmatrix} 3 & 8 & 8 & 8 \\ 8 & 3 & 8 & 8 \\ 8 & 8 & 3 & 8 \\ 8 & 8 & 8 & 3 \end{bmatrix} \quad \begin{bmatrix} 8 & 3 & 3 & 3 & 3 \\ 3 & 8 & 3 & 3 & 3 \\ 3 & 3 & 8 & 3 & 3 \\ 3 & 3 & 3 & 8 & 3 \\ 3 & 3 & 3 & 3 & 8 \end{bmatrix}$$

19. [M]用矩阵软件计算下列矩阵的行列式.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}$$

利用上述计算结果猜想下列矩阵的行列式，然后利用行变换计算行列式，以验证你的猜想.

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & n \end{bmatrix}$$

20. [M]利用习题 19 的方法猜想下列行列式：

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 3 & 3 & \cdots & 3 \\ 1 & 3 & 6 & \cdots & 6 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 3 & 6 & \cdots & 3(n-1) \end{bmatrix}$$

验证你的猜想. [提示：利用习题 14(c)和习题 19 的结果.] 213

第4章　向量空间

实例介绍：空间飞行与控制系统

1981年4月，在凉爽的棕枝全日[1]清晨，12层楼高、70吨重的哥伦比亚号航天飞机腾空而起. 这是美国历史上首架航天飞机，它可以说是控制系统工程设计领域历经十年深入研究和开发的成果，这一领域涉及工程学的许多分支——航空学、化学、电子学、流体力学和机械学.

对空间飞行来说，航天飞机的控制系统绝对是关键. 由于机身不稳定，航天飞机在大气飞行过程中需要计算机进行实时监控. 飞行监控系统向航天飞机的气动控制翼和44个小推进器发出一系列指令. 图4-1给出了一个典型的闭环反馈系统，它用于控制航天飞机在飞行中的倾角(即火箭头部锥体的上升角度). 计算机信号沿该图顶部传送，不同传感器的信号则从连接符号(⊗)标识的位置汇入其中.

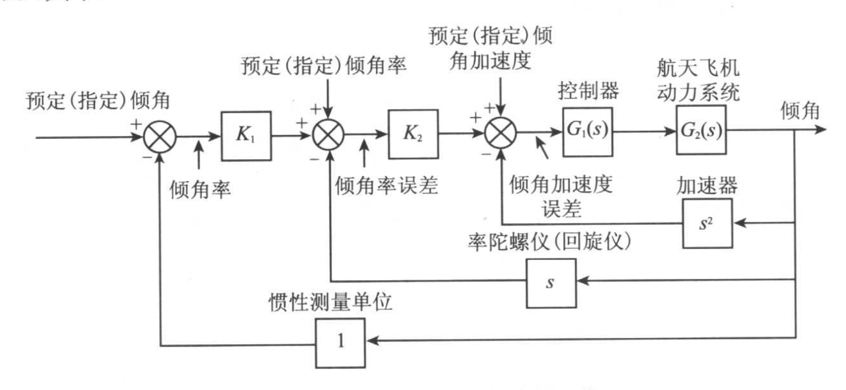


图4-1　航天飞机的倾角控制系统

(资料来源：*Control Systems Engineering*, by Norman S. Nise, Benjamin-Cummings Publishing, 1992, p. 274. 简图取自 *Space Shuttle GN& C Operations Manual*, Rockwell International, 1988.)

在数学意义上,工程系统中的输入输出信号都是函数. 这些函数可以如图4-1那样进行加法和数乘运算,这在实际应用中相当重要. 函数的这两种运算与我们将在4.4节和4.8节看到的$\mathbf{R}^n$中向量的加法和向量的数乘运算,有完全类似的代数性质. 有鉴于此,我们把所有可能的输入(函数)的集合称为向量空间. 而系统工程的数学基础就

---

1. 复活节前的星期日. ——译者注

依赖于这些函数所构成的向量空间. 后面, 我们还会看到其他向量空间怎样在工程学、物理学和统计学中出现.

215

我们在第 1 章和第 2 章播下的数学种子将会在这一章开花结果. 一旦认识到 $\mathbf{R}^n$ 仅仅是实际应用中产生的许多向量空间之一, 我们就能更清楚地体会到线性代数的力与美. 实际上, 研究向量空间和研究 $\mathbf{R}^n$ 本身没有太大的区别, 因为我们可以利用 $\mathbf{R}^2$ 和 $\mathbf{R}^3$ 上的几何经验使许多更一般的概念直观化.

从 4.1 节的基本定义开始, 一般向量空间的框架将在本章逐渐展开. 4.3 ~ 4.5 节的目标是说明其他的向量空间和 $\mathbf{R}^n$ 类似. 4.6 节关于秩的内容是本章的重点之一, 它利用向量空间的术语将矩阵的主要内容联系起来. 4.8 节把这一章的理论应用到航天飞行等数字控制系统的离散信号和差分方程中. 4.9 节的马尔可夫链, 一改本章理论性较强的风格, 为第 5 章将要引入的概念给出了很好的例子.

4.1 向量空间与子空间

第 1 章和第 2 章中的许多理论只依赖于 1.3 节列出的 $\mathbf{R}^n$ 上的某些浅显的代数性质. 实际上, 许多其他的数学系统也具有同样的性质. 下面的定义详细列出了我们感兴趣的一些性质.

216

4.1.1 向量空间

【定义】 设 V 是一个非空集合, 其中的元素称为向量. V 上定义了两种运算, 分别称为加法和(实数)数乘. 如果 V 满足下述十条公理, 就称 V 是一个**向量空间**(vector space).[1]对于 V 中的任意向量 $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ 和任意实数 c, d, 这些公理都必须成立:

1. $\boldsymbol{u}$, $\boldsymbol{v}$ 的和属于 V, 记为 $\boldsymbol{u}+\boldsymbol{v}$.
2. $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$.
3. $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$.
4. 存在零向量 $\mathbf{0}$ 在 V 中, 满足 $\boldsymbol{u}+\mathbf{0}=\boldsymbol{u}$.
5. 对 V 中任一向量 $\boldsymbol{u}$, 存在向量 $-\boldsymbol{u}\in V$, 满足 $\boldsymbol{u}+(-\boldsymbol{u})=\mathbf{0}$.
6. $\boldsymbol{u}$ 和 c 的数量乘积, 记为 $c\boldsymbol{u}$, 属于 V.
7. $c(\boldsymbol{u}+\boldsymbol{v})=c\boldsymbol{u}+c\boldsymbol{v}$.
8. $(c+d)\boldsymbol{u}=c\boldsymbol{u}+d\boldsymbol{u}$.
9. $c(d\boldsymbol{u})=(cd)\boldsymbol{u}$.
10. $1\boldsymbol{u}=\boldsymbol{u}$.

仅利用这些公理, 我们就可以证明公理 4 中的零向量是唯一的. 对 V 的任一向量 $\boldsymbol{u}$, 公理 5 中的向量 $-\boldsymbol{u}$, 称为 $\boldsymbol{u}$ 的**负向量**(negative), 也是唯一的. 见习题 25 和 26. 下列简单事实的证明要点也在习题中:

对任意 V 中的向量 $\boldsymbol{u}$ 以及数量 c,

$$0\boldsymbol{u}=\mathbf{0} \tag{1}$$

$$c\mathbf{0}=\mathbf{0} \tag{2}$$

1. 严格地说, V 是一个实向量空间. 对数量取复数的复向量空间本章中的所有理论同样成立. 第 5 章将出现这类向量空间, 在此之前, 所有的数量都假定为实数.

$$-\boldsymbol{u}=(-1)\boldsymbol{u} \tag{3}$$

【例题1】　$\mathbf{R}^n(n\geqslant 1)$是向量空间最主要的例子，对$\mathbf{R}^3$的几何直观印象将有助于理解本章中许多概念，使其形象化. ■

【例题2】　设V是三维空间中所有箭头(有向线段)构成的集合，其中，两个箭头相等当且仅当它们的长度相等且方向相同. 利用平行四边形法则(见1.3节)定义V上的加法；对V中的每个$\boldsymbol{v}$，定义$c\boldsymbol{v}$为一个箭头，其长度是$\boldsymbol{v}$的$|c|$倍，如果$c\geqslant 0$，$c\boldsymbol{v}$所指的方向与$\boldsymbol{v}$同向，否则与$\boldsymbol{v}$反向(见图4-2). 证明V是一个向量空间. 这个空间是物理问题中各种力的一个常用的模型. ■

217

解：V是从几何角度出发来定义的，其中用到了长度和方向的概念，不涉及xyz坐标系. 长度为零的箭头是个单点，它代表了零向量，$\boldsymbol{v}$的负向量是$(-1)\boldsymbol{v}$. 因此公理1、4、5、6、10显然成立，其余公理可以通过几何方法验证. 如图4-3和图4-4.

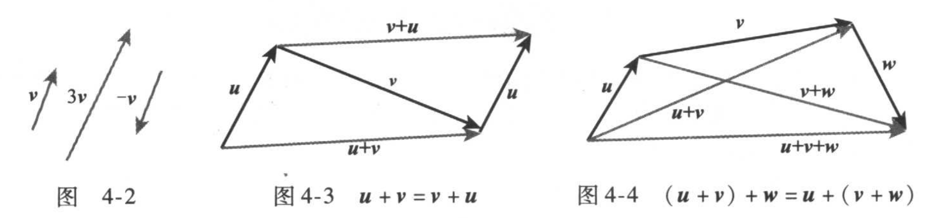


图　4-2　　图4-3　$\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$　　图4-4　$(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$

【例题3】　设$\mathbf{S}$是由两端都无限的全体数值序列所构成的空间(通常写成一行而非一列)：

$$\{y_k\}=(\cdots,y_{-2},y_{-1},y_0,y_1,y_2,\cdots)$$

如果$\{z_k\}$是$\mathbf{S}$中的另一个元素，$\{y_k\}+\{z_k\}$是$\{y_k\}$和$\{z_k\}$中的对应项相加得到的新序列$\{y_k+z_k\}$. 数量乘积$c\{y_k\}$是序列$\{cy_k\}$. 与$\mathbf{R}^n$类似，我们可以验证在$\mathbf{S}$中向量空间的公理也成立.

$\mathbf{S}$中的元素经常出现在工程学中. 例如，在离散的时间段上信号的测量(或采样). 这种信号可以是电子的、机械的、光学的等. 如本章实例介绍所述，航天飞机的中央控制系统使用了离散信号. 为使用方便，我们称$\boldsymbol{S}$为(离散时间)**信号**(signal)空间. 图4-5给出了它的直观含义. ■

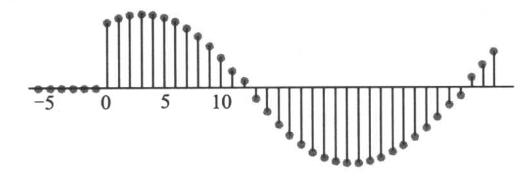


图4-5　离散时间信号

【例题4】　对$n\geqslant 0$，次数不超过n的多项式集合$\mathbf{P}_n$由如下形式的多项式构成：

$$\boldsymbol{p}(t)=a_0+a_1t+a_2t^2+\cdots+a_nt^n \tag{4}$$

其中$a_0,\cdots,a_n$和变量t是实数，多项式(4)中最高次项的系数不为零. 如果$p(t)=a_0\neq 0$，则$\boldsymbol{p}$的次数为零. 如果所有的系数都为零，则称$\boldsymbol{p}$为*零多项式*. 零多项式包含于$\mathbf{P}_n$，尽管没有定义它的次数.

218

如果 $\boldsymbol{p}$ 是如(4)给定，$\boldsymbol{q}(t)=b_0+b_1t+\cdots+b_nt^n$，则 $\boldsymbol{p}+\boldsymbol{q}$ 定义为

$$(\boldsymbol{p}+\boldsymbol{q})(t)=\boldsymbol{p}(t)+\boldsymbol{q}(t)=(a_0+b_0)+(a_1+b_1)t+\cdots+(a_n+b_n)t^n$$

数量乘积 $c\boldsymbol{p}$ 的多项式定义为

$$(c\boldsymbol{p})(t)=c\boldsymbol{p}(t)=ca_0+(ca_1)t+\cdots+(ca_n)t^n$$

这种定义满足公理1和6，因为 $\boldsymbol{p}+\boldsymbol{q}$ 和 $c\boldsymbol{p}$ 都是次数不大于 n 的多项式．公理2、3和7~10可以由实数的性质得到．显然零多项式即为公理4中的零向量．最后，$(-1)\boldsymbol{p}$ 是 $\boldsymbol{p}$ 的负向量，故公理5成立．因此 $\mathbf{P}_n$ 是一个向量空间．

这类向量空间在6.8节将要讨论的数据的统计趋势分析中会用到． ■

【例题5】 设 V 是定义在某个集合 $\mathbf{D}$（譬如，$\mathbf{D}$ 是实数集或是实数轴上的某区间）上的全体实值函数所构成的集合，用通常的方法定义 V 上的加法和数量乘积：函数 $\boldsymbol{f}+\boldsymbol{g}$ 在定义域 $\mathbf{D}$ 中某点 t 的函数值为 $\boldsymbol{f}(t)+\boldsymbol{g}(t)$．同样地，对数量 c 和 V 中的向量 $\boldsymbol{f}$，数量乘积 $c\boldsymbol{f}$ 在 t 点的值为 $c\boldsymbol{f}(t)$．例如，如果 $\mathbf{D}=\mathbf{R}$，$\boldsymbol{f}(t)=1+\sin 2t$，$\boldsymbol{g}(t)=2+0.5t$，则

$$(\boldsymbol{f}+\boldsymbol{g})(t)=3+\sin 2t+0.5t \quad 且 \quad (2\boldsymbol{g})(t)=4+t$$

V 中两个函数相等当且仅当它们在 $\mathbf{D}$ 内的每一点 t 的函数值都相同．因此 V 中的零向量是恒等于零的函数，即对任意 t 都有 $\boldsymbol{f}(t)=0$，此外 $\boldsymbol{f}$ 的负向量为 $(-1)\boldsymbol{f}$．公理1和6显然成立，由实数的性质可推得其他公理亦成立，所以 V 是一个向量空间． ■

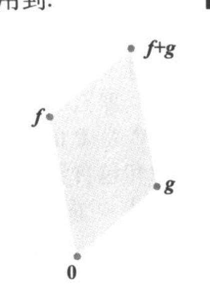


图4-6 两向量（函数）之和

将例题5向量空间 V 中的每一个函数看成单独的一个对象（比如向量空间中的"点"或者向量），这种观点非常重要．因为这可以帮助我们把在向量空间 $\mathbf{R}^n$ 中培养的几何直观推广到一般的向量空间，图4-6绘出了两个向量 $\boldsymbol{f}$ 与 $\boldsymbol{g}$（$\boldsymbol{f}$，$\boldsymbol{g}$ 可以是 V 中的函数或任意向量空间的元素）的和．参阅学习指南将有助于你理解和接受这种一般化的观点．

4.1.2 子空间

在许多问题中，所研究的向量空间往往由某个更大的向量空间的一个适当大小的子集所构成．在这种情况下，要说明该子集是向量空间，只需要验证其中的三条公理即可，其他公理自然成立．

219

【定义】 V 是一个向量空间，H 是 V 的子集，若 H 满足下列三条性质，则称 H 是 V 的子空间：

(a) V 的零向量属于 H.[1]

(b) H 中的向量满足加法封闭，即任意 $\boldsymbol{u}$，$\boldsymbol{v}\in H$，和向量 $\boldsymbol{u}+\boldsymbol{v}$ 属于 H.

(c) H 在数乘运算下封闭，即任意 $\boldsymbol{u}\in H$ 和数量 c，向量 $c\boldsymbol{v}$ 属于 H.

性质(a)，(b)和(c)保证了在 V 中已经定义的运算下，子空间 H 自身构成一个

1. 在有些教材中会通过假设 H 是非空的来代替性质(a)，则(a)可以由(c)和 $0\boldsymbol{u}=\mathbf{0}$ 导出．但是判定子空间最好的方法是，首先看零向量，如果 $\mathbf{0}$ 在 H 中，则需要验证性质(b)和(c)；如果 $\mathbf{0}$ 不在 H 中，则 H 不可能是子空间，其他性质就不需要验证了．

向量空间. 注意到性质(a), (b)和(c)就是公理1, 4和6. 公理2, 3和7~10在H中自然成立, 因为它们适用于V中的所有元素. 公理5在H中也成立, 因为(c)保证了如果$\boldsymbol{u}\in H$, 则$(-1)\boldsymbol{u}\in H$, 并且由等式(3)可知$(-1)\boldsymbol{u}$就是公理5中的向量$\boldsymbol{u}$.

所以, 每个子空间是一个向量空间. 反过来, 每个向量空间也是一个子空间(属于本身或其他更大的向量空间). 只有在涉及至少两个向量空间, 并且其中一个包含另一个时, 才用到子空间的概念. 当说到V的子空间时意思是把V看作是较大的空间. (见图4-7)

【例题6】 向量空间V中只含零向量的集合是V的一个子空间, 称为**零子空间**(zero subspace), 记作$\{\mathbf{0}\}$. ■

【例题7】 设$\mathbf{P}$是所有实系数的多项式组成的集合, $\mathbf{P}$中的运算与函数空间内的定义相同. 那么$\mathbf{P}$是所有定义在$\mathbf{R}$上的实值函数所构成空间的子空间. 并且, 对任意$n\geqslant 0$, $\mathbf{P}_n$是$\mathbf{P}$的一个子空间, 这是因为$\mathbf{P}_n$作为$\mathbf{P}$的子集含有零多项式, $\mathbf{P}_n$中两个多项式的和包含于$\mathbf{P}_n$, $\mathbf{P}_n$中多项式的数乘也包含于$\mathbf{P}_n$. ■

【例题8】 向量空间$\mathbf{R}^2$不是$\mathbf{R}^3$的子空间, 因为$\mathbf{R}^2$根本不是$\mathbf{R}^3$的子集($\mathbf{R}^3$中的向量有三个分量, 但$\mathbf{R}^2$中的分量却只有两个). 集合

$$H=\left\{\begin{bmatrix}s\\t\\0\end{bmatrix}: s\text{ 和 }t\text{ 是实数}\right\}$$

是$\mathbf{R}^3$的与$\mathbf{R}^2$有相同表现的子集, 尽管严格意义上H不同于$\mathbf{R}^2$. 见图4-8. 证明H是$\mathbf{R}^3$子空间.

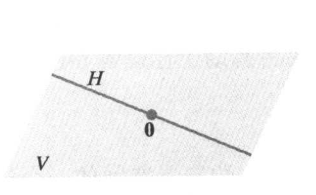


图4-7 V的子空间

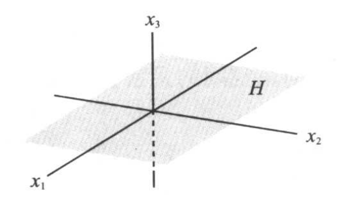


图4-8 将x_1x_2平面看成$\mathbf{R}^3$的子空间

解: 零向量属于H, 并且在向量的加法和数乘运算下H保持封闭, 因为在H中, 这两种运算所得新向量的第三个分量都是零(因而属于H), 因此H是$\mathbf{R}^3$的子空间. ■

220

【例题9】 在$\mathbf{R}^3$中不过原点的平面不是它的子空间, 因为该平面不包含$\mathbf{R}^3$中的零向量. 类似地, 在$\mathbf{R}^2$中不过原点的直线也不是它的子空间, 如图4-9所示. ■

4.1.3 集合张成的子空间

下面的例题给出了描述子空间的一种最普遍的方法. 与第1章相同, **线性组合**(linear combination)是指由向量的数量乘积之和构成的集合, Span$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$表示所有可以写成$\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p$的线性组合的向量所组成的集合.

【例题10】 假设$\boldsymbol{v}_1$, $\boldsymbol{v}_2$属于向量空间V, $H=\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$. 证明$H$是$V$的子空间.

解: 零向量属于H, 因为$\mathbf{0}=0\boldsymbol{v}_1+0\boldsymbol{v}_2$. 为证明$H$在向量的加法下封闭, 在$H$中任取两个向量,

$$u = s_1 v_1 + s_2 v_2 \text{ 和 } w = t_1 v_1 + t_2 v_2$$

根据公理2、3 和 8，对向量空间 V，

$$u + w = (s_1 v_1 + s_2 v_2) + (t_1 v_1 + t_2 v_2) = (s_1 + t_1) v_1 + (s_2 + t_2) v_2$$

所以，$u + w \in H$. 而且，如果 c 是任意数量，则由公理 7 和 9 可以得到

$$cu = c(s_1 v_1 + s_2 v_2) = (cs_1) v_1 + (cs_2) v_2$$

因此 $cu \in H$，H 在数乘运算下封闭. 所以 H 是 V 的子空间. ■

在 4.5 节，我们将会看到，$\mathbf{R}^3$ 的所有非零子空间，除了它本身之外，或者是 Span$\{v_1, v_2\}$(其中 v_1, v_2线性无关)，或者是 Span$\{v\}$($v \neq \mathbf{0}$). 在第一种情况下，子空间是过原点的平面；第二种情况下，子空间是过原点的直线(见图 4-10).

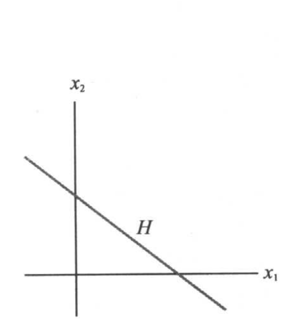


图 4-9 直线不是向量空间

图 4-10 一个子空间的例子

推广例题 10 中的思想可以证明下面的定理.

【定理 1】 如果 $v_1, \cdots, v_p$属于向量空间 V，则 Span$\{v_1, \cdots, v_p\}$是 V 的子空间.

我们把 Span$\{v_1, \cdots, v_p\}$称为**由**$\{v_1, \cdots, v_p\}$**张成的子空间**(the subspace spanned by $\{v_1, \cdots, v_p\}$)或**由**$\{v_1, \cdots, v_p\}$**生成的子空间**(the subspace generated by $\{v_1, \cdots, v_p\}$). 假设 H 是 V 的子空间，称$\{v_1, \cdots, v_p\}$是 H 的**张成集**(spanning set)或**生成集**(generating set)，如果 $H = \text{Span}\{v_1, \cdots, v_p\}$.

下面的例题是定理 1 的应用.

221

【例题 11】 设 H 是所有形如$(a-3b, b-a, a, b)$的向量所组成的集合，其中 a，b 是任意的数量. 即 $H = \{(a-3b, b-a, a, b) : a, b \in \mathbf{R}\}$. 证明 H 是 $\mathbf{R}^4$ 的子空间.

解： 把 H 中的向量记成列向量的形式，H 中的任意向量都具有如下的形式：

$$\begin{bmatrix} a-3b \\ b-a \\ a \\ b \end{bmatrix} = a \underbrace{\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}}_{v_1} + b \underbrace{\begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix}}_{v_2}$$

这表明 $H = \text{Span}\{v_1, v_2\}$，其中 v_1, v_2 如上所示. 因此由定理 1 知 H 是 $\mathbf{R}^4$ 的子空间. ■

例题 11 给出了一种非常有用的技巧，利用它可以将子空间 H 表示成某个更小的向

量集合的线性组合. 如果 $H=\mathrm{Span}\{v_1,\cdots,v_p\}$，我们可以利用其中的向量 $v_1,\cdots,v_p$ 来研究子空间 H. 经常把 H 中无穷多个向量的计算化简成张成集中有限多个向量的运算.

【例题 12】 如果

$$v_1=\begin{bmatrix}1\\-1\\-2\end{bmatrix},v_2=\begin{bmatrix}5\\-4\\-7\end{bmatrix},v_3=\begin{bmatrix}-3\\1\\0\end{bmatrix},\text{和 } y=\begin{bmatrix}-4\\3\\h\end{bmatrix}$$

则 h 取何值时，y 会属于由 v_1，v_2，v_3 张成的 $\mathbf{R}^3$ 的子空间.

解： 该问题是 1.3 节的基础练习 2，这里用子空间代替了 $\mathrm{Span}\{v_1,v_2,v_3\}$. 在那里得到的答案是 $y\in\mathrm{Span}\{v_1,v_2,v_3\}$ 当且仅当 $h=5$. 同时可以回顾一下 1.3 节习题 11 ~ 14 以及 17 ~ 21. ■

尽管这一章里的许多向量空间都是 $\mathbf{R}^n$ 的子空间，但这些抽象的理论同样适用于其他向量空间. 由函数构成的向量空间有着广泛的应用，在以后的学习中会进一步地接触到.

基础练习

1. 证明：在 $\mathbf{R}^2$ 中，所有形如 $(3s,2+5s)$ 组成的集合 H 不是线性空间，利用数乘运算封闭性（即找到 H 中的一个向量 u 和数量 c 使得 $cu\notin H$）.
2. 设 $W=\mathrm{Span}\{v_1,\cdots,v_p\}$，其中 $v_1,\cdots,v_p$ 属于向量空间 V，证明对于 $1\leqslant k\leqslant p$，有 $v_k\in W$. [提示：首先写出一个说明 v_1 包含于 W 的等式，然后将其推广到一般的情况]

222

习题 4.1

1. 设 V 是 xy 平面中的第一象限，即 $V=\left\{\begin{bmatrix}x\\y\end{bmatrix}:x\geqslant 0,\quad y\geqslant 0\right\}$
   a. 如果 u，$v\in V$，则 $u+v$ 是否属于 V? 为什么?
   b. 求 V 中一个特殊的向量 u 以及一个特殊的数量 c，使 $cu\notin V$.（这充分说明了 V 不是一个向量空间.）
2. 设 W 是 xy 平面中第一、三象限的并，即公式 $W=\left\{\begin{bmatrix}x\\y\end{bmatrix}:xy\geqslant 0\right\}$
   a. 如果 $u\in W$，c 是任意数量，则 cu 是否属于 W? 为什么?
   b. 求 W 中的两个向量 u，v，使 $u+v\notin W$. 这充分说明了 W 不是一个向量空间.
3. 设 H 是 xy 平面中由单位圆上及其内部的点组成的集合，即 $H=\left\{\begin{bmatrix}x\\y\end{bmatrix}:x^2+y^2\leqslant 1\right\}$ 找出一个特殊的例子(两个向量或者一个向量、一个数量)来说明 H 不是 $\mathbf{R}^2$ 的子空间.
4. 构造一个几何图形来说明为什么在 $\mathbf{R}^2$ 中不过原点的直线在向量的加法运算下不封闭.

在习题 5 ~ 8 中，判断对于适当的 n，所给定的集合是否是 $\mathbf{R}^n$ 的子空间，并说明理由.

5. 所有形如 $p(t)=at^2$ 的多项式，其中 $a\in\mathbf{R}$.
6. 所有形如 $p(t)=a+t^2$ 的多项式，其中 $a\in\mathbf{R}$.
7. 所有次数至多为 3 的整系数多项式.

8. 在 $\mathbf{P}_n$ 中所有满足 $\boldsymbol{p}(0)=0$ 的多项式.

9. 设 H 是所有形如 $\begin{bmatrix} s \\ 3s \\ 2s \end{bmatrix}$ 的向量组成的集合. 试在 $\mathbf{R}^3$ 中找到一个向量 $\boldsymbol{v}$ 满足 $H=\text{Span}\{\boldsymbol{v}\}$. 为什么仅此就可以说明 H 是 $\mathbf{R}^3$ 的一个子空间?

10. 设 H 是所有形如 $\begin{bmatrix} 2t \\ 0 \\ -t \end{bmatrix}$ 的向量组成的集合，证明：H 是 $\mathbf{R}^3$ 的子空间(利用习题 9 的方法).

11. 设 W 是所有形如 $\begin{bmatrix} 5b+2c \\ b \\ c \end{bmatrix}$ 的向量组成的集合，其中 b, c 任意，求向量 $\boldsymbol{u}$, $\boldsymbol{v}$, 使 $W=\text{Span}\{\boldsymbol{u},\boldsymbol{v}\}$. 为什么仅此就可以说明 H 是 $\mathbf{R}^3$ 的一个子空间?

12. 设 W 是所有形如 $\begin{bmatrix} s+3t \\ s-t \\ 2s-t \\ 4t \end{bmatrix}$ 的向量组成的集合. 证明：W 是 $\mathbf{R}^4$ 的一个子空间(利用习题 11 中的方法).

13. 设 $\boldsymbol{v}_1=\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \boldsymbol{v}_2=\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \boldsymbol{v}_3=\begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix}$, 且 $\boldsymbol{w}=\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$.

a. $\boldsymbol{w}$ 是否包含于 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$? $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 中包含多少向量?
b. $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ 包含多少向量?
c. $\boldsymbol{w}$ 是否包含于由 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ 张成的子空间内? 为什么?

14. 设 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 同上题，$\boldsymbol{w}=\begin{bmatrix} 8 \\ 4 \\ 7 \end{bmatrix}$. 则 $\boldsymbol{w}$ 是否属于由 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 张成的子空间? 为什么?

在习题 15 ~ 18 中，设 W 是由所有形如题中所给形式的向量构成的集合，a, b, c 是任意实数，在每一种情况下，或者求出 W 的一个张成集，或者找出一个例子说明 W 不是线性空间.

15. $\begin{bmatrix} 3a+b \\ 4 \\ a-5b \end{bmatrix}$　16. $\begin{bmatrix} -a+1 \\ a-6b \\ 2b+a \end{bmatrix}$　17. $\begin{bmatrix} a-b \\ b-c \\ c-a \\ b \end{bmatrix}$　18. $\begin{bmatrix} 4a+3b \\ 0 \\ a+b+c \\ c-2a \end{bmatrix}$

19. 将一质量为 m 的物体挂在弹簧末端，向下拉动物体，然后释放，则弹簧振子系统将开始振荡. 物体离开初始位置的位移 y 可由以下函数给出：

$$y(t)=c_1\cos\omega t+c_2\sin\omega t \quad (5)$$

其中 ω 是依赖于弹簧和物体的常数(见右图). 证明：(5)描述的所有函数(ω 固定, c_1, c_2 任意)构成一个线性空间.

y

习题 19 的图

20. $C[a,b]$ 表示定义在 $\mathbf{R}$ 上闭区间 $[a,b]$ 内的所有连续的实值函数所构成的集合. 该集合是 $[a,b]$ 上所有实值函数构成的向量空间的子空间.

a. 为说明 $C[a,b]$ 确实是一个子空间，应该证明连续函数的哪些性质？(这些性质常在微积分中涉及到).

b. 证明：$\{f\in C[a,b]: f(a)=f(b)\}$ 是 $C[a,b]$ 的一个子空间.

对固定的正整数 m 和 n，由全体 $m\times n$ 矩阵所构成的集合 $M_{m\times n}$ 在通常的矩阵加法和数乘运算下构成向量空间.

21. 试判断由全体形如 $\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ 的矩阵所构成的集合 H 是否是 $M_{2\times 2}$ 的子空间.

22. 设 $\boldsymbol{F}$ 是一个固定的 3×2 矩阵，H 是 $M_{2\times 4}$ 中全体满足 $\boldsymbol{FA}=\boldsymbol{0}$ ($M_{3\times 4}$ 中的零矩阵) 的矩阵 $\boldsymbol{A}$ 所构成的集合，试判定 H 是否是 $M_{2\times 4}$ 的子空间.

在习题 23 和 24 中，判断命题的真假，并说明理由.

23. a. 设 V 是由实数域上全体实值函数构成的向量空间，如果 $f\in V$ 并且满足对某个 t，$f(t)=0$，那么 f 是 V 中的零向量.

b. 在三维空间中，向量是一个箭头.

c. 设 H 是向量空间 V 的一个子集，如果 V 的零向量属于 H，则 H 是 V 的子空间.

d. 子空间也是向量空间.

e. 如本章实例介绍所述，航天飞机的中央控制系统中使用了模拟信号.

24. a. 向量空间中每个元素都是向量.

b. 如果 $\boldsymbol{u}$ 是向量空间 V 中的向量，则 $(-1)\boldsymbol{u}$ 就是 $\boldsymbol{u}$ 的负向量.

c. 向量空间也是一个子空间.

d. $\mathbf{R}^2$ 是 $\mathbf{R}^3$ 的一个子空间.

e. 如果满足下列条件，那么向量空间 V 的子集 H 是 V 的子空间.

(i) V 的零向量属于 H,

(ii) $\boldsymbol{u}$，$\boldsymbol{v}$ 和 $\boldsymbol{u}+\boldsymbol{v}$ 属于 H,

(iii) c 是一数量，并且 $c\boldsymbol{u}\in H$.

习题 25～29 演示了如何利用向量空间的公理来证明向量空间的基本性质，试在空格中填上恰当的公理序号. 根据公理 2、公理 4 和公理 5 分别蕴含：对任意 $\boldsymbol{u}$，有 $\boldsymbol{0}+\boldsymbol{u}=\boldsymbol{u}$ 以及 $-\boldsymbol{u}+\boldsymbol{u}=\boldsymbol{0}$.

25. 完成下面关于零向量唯一性的证明. 假设 $\boldsymbol{w}\in V$，并且满足 $\boldsymbol{u}+\boldsymbol{w}=\boldsymbol{w}+\boldsymbol{u}=\boldsymbol{u}$ 对所有 V 中的 $\boldsymbol{u}$ 都成立，特别地有，$\boldsymbol{0}+\boldsymbol{w}=\boldsymbol{w}$，由公理______，所以 $\boldsymbol{w}=\boldsymbol{0}+\boldsymbol{w}=\boldsymbol{0}$.

26. 完成下面关于对于任意 $\boldsymbol{u}\in V$，满足 $\boldsymbol{u}+(-\boldsymbol{u})=\boldsymbol{0}$ 的负向量 $(-\boldsymbol{u})$ 的唯一性的证明. 假设 $\boldsymbol{w}$ 满足 $\boldsymbol{u}+\boldsymbol{w}=\boldsymbol{0}$，等式两边同时加 $(-\boldsymbol{u})$，我们有

$(-\boldsymbol{u})+[\boldsymbol{u}+\boldsymbol{w}]=(-\boldsymbol{u})+\boldsymbol{0}$

$[(-\boldsymbol{u})+\boldsymbol{u}]+\boldsymbol{w}=(-\boldsymbol{u})+\boldsymbol{0}$　　由公理______ (a)

$\boldsymbol{0}+\boldsymbol{w}=(-\boldsymbol{u})+\boldsymbol{0}$　　由公理______ (b)

$\boldsymbol{w}=-\boldsymbol{u}$　　由公理______ (c)

27. 在下面关于对每一个 $\boldsymbol{u}\in V$ 都满足 $0\boldsymbol{u}=\boldsymbol{0}$ 的证明中填上所用到的公理的序号.

$0\boldsymbol{u}=(0+0)\boldsymbol{u}=0\boldsymbol{u}+0\boldsymbol{u}$　　由公理______ (a)

在等式的两边同时加 $0\boldsymbol{u}$ 的负向量：

$0\boldsymbol{u}+(-0\boldsymbol{u})=[0\boldsymbol{u}+0\boldsymbol{u}]+(-0\boldsymbol{u})$

$0\boldsymbol{u}+(-0\boldsymbol{u})=0\boldsymbol{u}+[0\boldsymbol{u}+(-0\boldsymbol{u})]$　　由公理______ (b)

$\boldsymbol{0}=0\boldsymbol{u}+\boldsymbol{0}$　　由公理______ (c)

$\boldsymbol{0}=0\boldsymbol{u}$　　由公理______ (d)

28. 在下面关于对任意数量 c 都有 $c\boldsymbol{0}=\boldsymbol{0}$ 的证明中填上所用到的公理的序号.

$c\mathbf{0} = c(\mathbf{0}+\mathbf{0})$ 由公理______(a)

$= c\mathbf{0} + c\mathbf{0}$ 由公理______(b)

在等式的两边同时加 $c\mathbf{0}$ 的负向量

$c\mathbf{0} + (-c\mathbf{0}) = [c\mathbf{0} + c\mathbf{0}] + (-c\mathbf{0})$

$c\mathbf{0} + (-c\mathbf{0}) = c\mathbf{0} + [c\mathbf{0} + (-c\mathbf{0})]$ 由公理______(c)

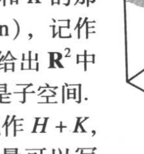
$\mathbf{0} = c\mathbf{0} + \mathbf{0}$ 由公理______(d)

$\mathbf{0} = c\mathbf{0}$ 由公理______(e)

29. 证明 $(-1)\boldsymbol{u} = -\boldsymbol{u}$. [提示：证明 $\boldsymbol{u} + (-1)\boldsymbol{u} = \mathbf{0}$，利用公理和习题 27 和 26 的结果.] 224

30. 假设对某个非零数量 c 有 $c\boldsymbol{u} = \mathbf{0}$，证明 $\boldsymbol{u} = \mathbf{0}$，并写出你所用到的公理或性质.

31. 设 $\boldsymbol{u}$，$\boldsymbol{v}$ 包含于向量空间 V，H 是 V 的任一包含 $\boldsymbol{u}$，$\boldsymbol{v}$ 的子空间. 解释 H 为什么也包含 $\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$. 这说明 $\mathrm{Span}\{\boldsymbol{u},\boldsymbol{v}\}$ 是 V 的最小的同时包含 $\boldsymbol{u}$，$\boldsymbol{v}$ 的子空间.



32. 设 H 和 K 是向量空间 V 的子空间. V 中同时包含于 H 和 K 的那些元素所组成的集合，称为 H 和 K 的**交集**(intersection)，记作 $H \cap K$. 证明：$H \cap K$ 是 V 的一个子空间(见右图). 试给出 $\mathbf{R}^2$ 中的一个例子来说明在一般情况下两个子空间的并集不是子空间.

33. 假设 H 和 K 是向量空间 V 的子空间，H 和 K 的和，记作 $H+K$，是 V 中所有满足下面条件的向量所组成的集合，此向量可以写成两个向量的和，其中一个属于 H，另一个属于 K，即

$H + K = \{\boldsymbol{w} : \boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v} \text{ 其中 } \boldsymbol{u} \in H, \boldsymbol{v} \in K\}$

a. 证明：$H+K$ 是 V 的子空间.

b. 证明：H 是 $H+K$ 的子空间并且 K 是 $H+K$ 的子空间.

34. 假设 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p$ 和 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_q$ 是向量空间 V 中的向量，并且设

$$H = \mathrm{Span}\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p\}, K = \mathrm{Span}\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_q\}$$

证明：$H + K = \mathrm{Span}\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p, \boldsymbol{v}_1 \cdots, \boldsymbol{v}_q\}$

35. [**M**]证明：$\boldsymbol{w}$ 属于由 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ 所张成的 $\mathbf{R}^4$ 的子空间，其中

$$\boldsymbol{w} = \begin{bmatrix} -9 \\ 7 \\ 4 \\ 8 \end{bmatrix}, \quad \boldsymbol{v}_1 = \begin{bmatrix} 7 \\ -4 \\ -2 \\ 9 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} -4 \\ 5 \\ -1 \\ -7 \end{bmatrix}, \quad \boldsymbol{v}_3 = \begin{bmatrix} -9 \\ 4 \\ 4 \\ -7 \end{bmatrix}$$

36. [**M**]判定 $\boldsymbol{y}$ 是否属于由 $\boldsymbol{A}$ 的列向量张成的子空间，其中

$$\boldsymbol{y} = \begin{bmatrix} 6 \\ 7 \\ 1 \\ -4 \end{bmatrix}, \quad \boldsymbol{A} = \begin{bmatrix} 5 & -5 & -9 \\ 8 & 8 & -6 \\ -5 & -9 & 3 \\ 3 & -2 & -7 \end{bmatrix}$$

37. [**M**]向量空间 $H = \mathrm{Span}\{1, \cos^2 t, \cos^4 t, \cos^6 t\}$ 至少包含下道习题中将用到的两个有趣的函数：

$$\boldsymbol{f}(t) = 1 - 8\cos^2 t + 8\cos^4 t$$
$$\boldsymbol{g}(t) = -1 + 18\cos^2 t - 48\cos^4 t + 32\cos^6 t$$

研究函数 $\boldsymbol{f}$ 在 $0 \leqslant t \leqslant 2\pi$ 上的图像，并且猜测 $\boldsymbol{f}(t)$ 的一个简洁公式. 绘图比较 $1+\boldsymbol{f}(t)$ 与你所猜测的 $\boldsymbol{f}(t)$，以验证你的猜测(预期的结果是对任意 t，两个图像上的取值相减都得 1). 对 $\boldsymbol{g}$ 重复上述过程.

38. [**M**]在向量空间 $\mathrm{Span}\{1, \sin t, \sin^2 t, \cdots \sin^5 t\}$ 中，对下列函数重做习题 37.

$$f(t)=3\sin t-4\sin^3 t$$
$$g(t)=1-8\sin^2 t+8\sin^4 t$$
$$h(t)=5\sin t-20\sin^3 t+16\sin^5 t$$

基础练习答案

1. 取任意向量 $\boldsymbol{u}\in H$，比如说 $\boldsymbol{u}=\begin{bmatrix}3\\7\end{bmatrix}$，取任意 $c\neq 1$，比如说 $c=2$．则 $c\boldsymbol{u}=\begin{bmatrix}6\\14\end{bmatrix}$.

   如果 $c\boldsymbol{u}\in H$，则存在某个 s 满足

$$\begin{bmatrix}3s\\2+5s\end{bmatrix}=\begin{bmatrix}6\\14\end{bmatrix}$$

225

   也就是，$s=2$，且 $s=12/5$，这是不可能的．所以 $2\boldsymbol{u}\notin H$，H 不是一个向量空间.

2. $\boldsymbol{v}_1=1\boldsymbol{v}_1+0\boldsymbol{v}_2+\cdots+0\boldsymbol{v}_p$．即 $\boldsymbol{v}_1$ 是 $\boldsymbol{v}_1,\cdots\cdots,\boldsymbol{v}_p$ 的线性组合，所以 $\boldsymbol{v}_1\in W$．一般地，$\boldsymbol{v}_k\in W$，因为

$$\boldsymbol{v}_k=0\boldsymbol{v}_1+\cdots+0\boldsymbol{v}_{k-1}+1\boldsymbol{v}_k+0\boldsymbol{v}_{k+1}+\cdots+0\boldsymbol{v}_p$$

4.2　零空间、列空间和线性变换

在线性代数的应用中，$\mathbf{R}^n$的子空间通常以下列两种形式出现：(1)作为齐次线性方程组的所有解的集合，(2)作为某些特定向量的全体线性组合的集合．在本节中，我们反复利用子空间的概念，以比较和对照这两种子空间的描述．实际上，你很快会发现，从1.3节以来，我们一直在用子空间．这里主要的特点是更为专业化．本节最后是关于线性变换的核和值域的讨论.

4.2.1　矩阵的零空间

考虑下列齐次线性方程组

$$\begin{aligned}x_1-3x_2-2x_3&=0\\-5x_1+9x_2+x_3&=0\end{aligned}\tag{1}$$

如果用矩阵的形式来表示，该方程组可以写成 $A\boldsymbol{x}=\mathbf{0}$，其中

$$A=\begin{bmatrix}1&-3&-2\\-5&9&1\end{bmatrix}\tag{2}$$

回忆所有满足(1)的 $\boldsymbol{x}$ 组成的集合被称为方程组(1)的**解集**(solution set)．通常我们可以十分方便地将这个集合直接与矩阵 A 和方程 $A\boldsymbol{x}=\mathbf{0}$ 联系起来．我们称满足 $A\boldsymbol{x}=\mathbf{0}$ 的 $\boldsymbol{x}$ 组成的集合为矩阵 A 的**零空间**(null space).

【定义】　$m\times n$ 矩阵 A 的**零空间**(null space)，记作 Nul A，是齐次方程 $A\boldsymbol{x}=\mathbf{0}$ 的全体解组成的集合．用集合的符号表示如下：

$$\text{Nul } A=\{\boldsymbol{x}:\boldsymbol{x}\in\mathbf{R}^n, A\boldsymbol{x}=\mathbf{0}\}$$

Nul A 的一种更加动态的描述是：Nul A 是 $\mathbf{R}^n$中经线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 映至 $\mathbf{R}^m$中零向量的全体 $\boldsymbol{x}$ 所构成的集合(见图4-11).

226

【例题1】　设 A 和(2)式中相同，设 $\boldsymbol{u}=\begin{bmatrix}5\\3\\-2\end{bmatrix}$．判断 $\boldsymbol{u}$ 是否属于 A 的零空间.

解：为检验 $\boldsymbol{u}$ 是否满足 $A\boldsymbol{u}=\boldsymbol{0}$，计算

$$A\boldsymbol{u}=\begin{bmatrix}1 & -3 & -2\\ -5 & 9 & 1\end{bmatrix}\begin{bmatrix}5\\ 3\\ -2\end{bmatrix}=\begin{bmatrix}5-9+4\\ -25+27-2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

因此 $\boldsymbol{u}\in \mathrm{Nul}\,A$. ■

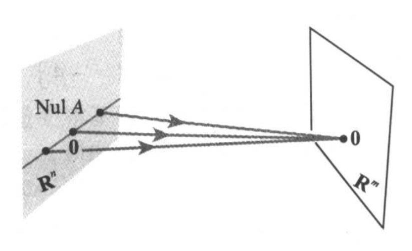


图　4-11

"零空间"中的"空间"一词十分恰当，因为一个矩阵的零空间就是一个向量空间，正如下面定理所述.

【定理2】　$m\times n$ 矩阵 A 的零空间是 $\mathbf{R}^n$ 的子空间. 等价地，含有 n 个未知量 m 齐次线性方程的方程组 $A\boldsymbol{x}=\boldsymbol{0}$ 的全体解构成的集合是 $\mathbf{R}^n$ 的子空间.

证明：显然 $\mathrm{Nul}\,A$ 是 $\mathbf{R}^n$ 的一个子集，因为 A 中有 n 列. 我们必须证明 $\mathrm{Nul}\,A$ 满足子空间的三条性质. 显然，$\boldsymbol{0}\in\mathrm{Nul}\,A$. 其次，设 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 代表 $\mathrm{Nul}\,A$ 中的任意两个向量，则

$$A\boldsymbol{u}=\boldsymbol{0}\text{ 且 }A\boldsymbol{v}=\boldsymbol{0}$$

为证明 $\boldsymbol{u}+\boldsymbol{v}\in\mathrm{Nul}\,A$，我们必须证明 $A(\boldsymbol{u}+\boldsymbol{v})=\boldsymbol{0}$. 由矩阵乘法的性质，我们知道

$$A(\boldsymbol{u}+\boldsymbol{v})=A\boldsymbol{u}+A\boldsymbol{v}=\boldsymbol{0}+\boldsymbol{0}=\boldsymbol{0}$$

因此 $\boldsymbol{u}+\boldsymbol{v}\in\mathrm{Nul}\,A$，并且 $\mathrm{Nul}\,A$ 在向量加法下封闭.

最后，如果 c 是任意数量，则

$$A(c\boldsymbol{u})=c(A\boldsymbol{u})=c(\boldsymbol{0})=\boldsymbol{0}$$

这说明 $c\boldsymbol{u}\in\mathrm{Nul}\,A$. 因此 $\mathrm{Nul}\,A$ 是 $\mathbf{R}^n$ 的子空间. ■

【例题2】　设 H 是 $\mathbf{R}^4$ 中满足 $a-2b+5c=d$ 和 $c-a=b$ 的全体向量 (a,b,c,d) 所构成的集合. 证明 H 是 $\mathbf{R}^4$ 的子空间.

227

解：通过整理描述 H 中元素性质的方程，我们可以看到 H 是下面齐次线性方程组的解集：

$$\begin{aligned}a-2b+5c-d&=0\\ -a-b+c&=0\end{aligned}$$

由定理2知，H 是 $\mathbf{R}^4$ 的子空间. ■

注意，在该例题中定义 H 的线性方程组是齐次的非常重要，否则，解集将不是子空间(因为零向量不是非齐次方程组的解). 而且，在有些情况下，解集有可能是空集.

4.2.2　Nul *A* 的显式表示

$\mathrm{Nul}\,A$ 中的向量与 A 中的向量之间没有明显的联系. 我们称 $\mathrm{Nul}\,A$ 是隐式定义

的，因为它是由一个待验证的条件定义的. 对 Nul $\boldsymbol{A}$ 中的元素并没有给出一个显式描述. 但是，求解方程 $\boldsymbol{Ax}=\boldsymbol{0}$ 将会给出 Nul $\boldsymbol{A}$ 的显式表示. 下面的例题复习了1.5节的算法.

【例题3】 求矩阵 $\boldsymbol{A}$ 的零空间的一个张成集.

$$\boldsymbol{A}=\begin{bmatrix}-3 & 6 & -1 & 1 & -7\\ 1 & -2 & 2 & 3 & -1\\ 2 & -4 & 5 & 8 & -4\end{bmatrix}$$

解: 首先用自由变量表示出 $\boldsymbol{Ax}=\boldsymbol{0}$ 的通解. 为了将基本变量写成自由变量的表达式，我们对增广矩阵 $[\boldsymbol{A}\quad \boldsymbol{0}]$ 进行行化简，得到简化阶梯矩阵:

$$\begin{bmatrix}1 & -2 & 0 & -1 & 3 & 0\\ 0 & 0 & 1 & 2 & -2 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\end{bmatrix},\qquad \begin{aligned}x_1-2x_2-x_4+3x_5&=0\\ x_3+2x_4-2x_5&=0\\ 0&=0\end{aligned}$$

求得通解为: $x_1=2x_2+x_4-3x_5$, $x_3=-2x_4+2x_5$, 其中 x_2, x_4 和 x_5 是自由变量.
下一步将表示通解的向量分解成以自由变量为权重的向量的线性组合，即

$$\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix}=\begin{bmatrix}2x_2+x_4-3x_5\\x_2\\-2x_4+2x_5\\x_4\\x_5\end{bmatrix}=x_2\underset{\boldsymbol{u}}{\begin{bmatrix}2\\1\\0\\0\\0\end{bmatrix}}+x_4\underset{\boldsymbol{v}}{\begin{bmatrix}1\\0\\-2\\1\\0\end{bmatrix}}+x_5\underset{\boldsymbol{w}}{\begin{bmatrix}-3\\0\\2\\0\\1\end{bmatrix}}\tag{3}$$

$$=x_2\boldsymbol{u}+x_4\boldsymbol{v}+x_5\boldsymbol{w}$$

向量 $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ 的每个线性组合都是 Nul $\boldsymbol{A}$ 中的元素，因此 $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ 是 Nul $\boldsymbol{A}$ 的张成集. ■

228

例题3的方法可以应用到解该类型的所有问题，这里需要指出两点:

1. 利用例题3的方法所求出的张成集是线性无关的，因为自由变量是张成向量的权重. 例如，观察(3)中解向量的第2、4、5个分量，注意到 $x_2\boldsymbol{u}+x_4\boldsymbol{v}+x_5\boldsymbol{w}$ 为 $\boldsymbol{0}$ 当且仅当权重 x_2, x_4 和 x_5 全部为零.

2. 如果 Nul $\boldsymbol{A}$ 中有非零向量，则 Nul $\boldsymbol{A}$ 的张成集中的向量个数等于方程 $\boldsymbol{Ax}=\boldsymbol{0}$ 中自由变量的个数.

4.2.3 矩阵的列空间

与矩阵相关的另一个重要子空间是矩阵的列空间. 与零空间不同的是，列空间是通过线性组合来显式定义的.

【定义】 $m\times n$ 矩阵 $\boldsymbol{A}$ 的**列空间**(column space)，记作 Col $\boldsymbol{A}$，是 $\boldsymbol{A}$ 的所有列向量的线性组合组成的集合. 若 $\boldsymbol{A}=[\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n]$，则

$$\text{Col}\,\boldsymbol{A}=\text{Span}\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n\}$$

因为 $\text{Span}\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n\}$ 是一个子空间，由定理1，从 Col $\boldsymbol{A}$ 的定义和 $\boldsymbol{A}$ 的列向量都属于 $\mathbf{R}^m$ 这一事实可以得到下面的定理.

【定理 3】 $m\times n$ 矩阵 $\boldsymbol{A}$ 的列空间是 $\mathbf{R}^m$ 的一个子空间.

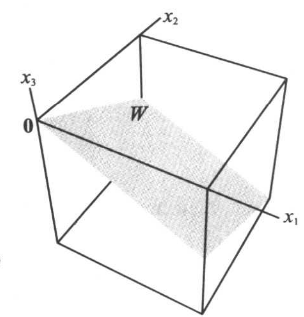
注意，Col $\boldsymbol{A}$ 中的向量可以记作 $\boldsymbol{Ax}$，这里 $\boldsymbol{x}$ 是某个向量. 这是因为 $\boldsymbol{Ax}$ 代表了 $\boldsymbol{A}$ 中列向量的一个线性组合，即

$$\boxed{\text{Col}\,\boldsymbol{A}=\{\boldsymbol{b}:\boldsymbol{b}=\boldsymbol{Ax},\boldsymbol{x}\in\mathbf{R}^n\}}$$

记号 $\boldsymbol{Ax}$ 也表明了 Col $\boldsymbol{A}$ 是线性变换 $\boldsymbol{x}\mapsto\boldsymbol{Ax}$ 的值域. 在本节末尾我们将继续讨论这一点.

【例题 4】 求使得 $W=\text{Col}\,\boldsymbol{A}$ 的矩阵 $\boldsymbol{A}$，其中

$$W=\left\{\begin{bmatrix}6a-b\\a+b\\-7a\end{bmatrix}:a,b\in\mathbf{R}\right\}$$



解： 首先，把 W 写成一个线性组合的集合.

$$W=\left\{a=\begin{bmatrix}6\\1\\-7\end{bmatrix}+b\begin{bmatrix}-1\\1\\0\end{bmatrix}:a,b\in\mathbf{R}\right\}=\text{Span}\left\{\begin{bmatrix}6\\1\\-7\end{bmatrix},\begin{bmatrix}-1\\1\\0\end{bmatrix}\right\}$$

229

接下来，将张成集中的向量作为矩阵 $\boldsymbol{A}$ 的列向量. 令 $\boldsymbol{A}=\begin{bmatrix}6&-1\\1&1\\-7&0\end{bmatrix}$，则 $W=\text{Col}\,\boldsymbol{A}$. ■

回顾 1.4 节定理 4，$\boldsymbol{A}$ 的列向量可以张成 $\mathbf{R}^m$ 当且仅当方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 对任意 $\boldsymbol{b}$ 都有解. 我们可以重述该事实为：

$m\times n$ 矩阵 $\boldsymbol{A}$ 的列空间是 $\mathbf{R}^m$ 当且仅当方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 对任意 $\boldsymbol{b}$ 都有解.

4.2.4 Nul A 和 Col A 的异同

学习了零空间和列空间之后，我们自然会思考一个问题：一个矩阵的零空间和列空间之间有何种联系？事实上，例题 5 ~ 7 将表明这两个空间有很大差异. 不过，零空间和列空间之间的确有着惊人的联系，这一点我们将在 4.6 节看到.

【例题 5】 设

$$\boldsymbol{A}=\begin{bmatrix}2&4&-2&1\\-2&-5&7&3\\3&7&-8&6\end{bmatrix}$$

a. 如果 $\boldsymbol{A}$ 的列空间是 $\mathbf{R}^k$ 的子空间，求 k 是多少？

b. 如果 $\boldsymbol{A}$ 的零空间是 $\mathbf{R}^k$ 的子空间，求 k 是多少？

解：

a. $\boldsymbol{A}$ 的每个列向量都含有三个分量，所以 Col $\boldsymbol{A}$ 是 $\mathbf{R}^k$ 的子空间，其中 $k=3$.

b. 使得 $\boldsymbol{Ax}$ 有意义的向量 $\boldsymbol{x}$ 必须含有四个分量，所以 Nul $\boldsymbol{A}$ 是 $\mathbf{R}^k$ 的子空间，其中 $k=4$. ■

当矩阵不是方阵时，比如在例题5中，Nul A 和 Col A 中的向量完全属于不同的"论域". 例如，$\mathbf{R}^3$ 中的向量的线性组合不可能是 $\mathbf{R}^4$ 中的一个向量. 如果 A 是方阵，则 Nul A 和 Col A 一定有共同的零向量，并且在特殊情况下，某些非零向量可能同时属于 Nul A 和 Col A.

230

【例题6】　设 A 如例题5，试分别找出 Col A 和 Nul A 中的一个非零向量.

解：找出 Col A 中的一个向量是非常容易的，A 中的任意列向量均可，比如取 $\begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}$. 为了找出 Nul A 中的一个非零向量，我们需要做些准备. 将增广矩阵 $[A\quad \mathbf{0}]$ 行化简可以得到

$$[A\quad \mathbf{0}] \sim \begin{bmatrix} 1 & 0 & 9 & 0 & 0 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

因此，如果 $\boldsymbol{x}$ 满足 $A\boldsymbol{x}=\mathbf{0}$，则 $x_1=-9x_3, x_2=5x_3, x_4=0$，且 x_3 是自由变量. 赋给 x_3 一个非零值，比如 $x_3=1$，我们可以得到 Nul A 中的一个向量，即 $\boldsymbol{x}=(-9,5,1,0)$. ■

【例题7】　设 A 同例题5，令 $\boldsymbol{u}=\begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix}, \boldsymbol{v}=\begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$.

a. 判断 $\boldsymbol{u}\in$ Nul A 是否成立，$\boldsymbol{u}$ 可能在 Col A 中吗？

b. 判断 $\boldsymbol{v}\in$ Col A 是否成立，$\boldsymbol{v}$ 可能在 Nul A 中吗？

解：

a. 这里不需要 Nul A 的显式表示，直接计算 $A\boldsymbol{u}$ 即可.

$$A\boldsymbol{u}=\begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix}=\begin{bmatrix} 0 \\ -3 \\ 3 \end{bmatrix}\neq\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

显然，$\boldsymbol{u}$ 不是 $A\boldsymbol{x}=\mathbf{0}$ 的解，所以 $\boldsymbol{u}\notin$ Nul A. 此外，由于 $\boldsymbol{u}$ 有四个分量，而 Col A 是 $\mathbf{R}^3$ 的子空间，所以 $\boldsymbol{u}$ 不可能在 Col A 中.

b. 将 $[A\quad \boldsymbol{v}]$ 化简成阶梯矩阵.

$$[A\quad \boldsymbol{v}]=\begin{bmatrix} 2 & 4 & -2 & 1 & 3 \\ -2 & -5 & 7 & 3 & -1 \\ 3 & 7 & -8 & 6 & 3 \end{bmatrix}\sim\begin{bmatrix} 2 & 4 & -2 & 1 & 3 \\ 0 & 1 & -5 & -4 & -2 \\ 0 & 0 & 0 & 17 & 1 \end{bmatrix}$$

显然方程 $A\boldsymbol{x}=\boldsymbol{v}$ 是相容的，所以 $\boldsymbol{v}\in$ Col A. 由于 $\boldsymbol{v}$ 只含三个分量，而 Nul A 是 $\mathbf{R}^4$ 的子空间，所以 $\boldsymbol{v}$ 不可能在 Nul A 中. ■

表4-1总结了我们已经学过的 Nul A 和 Col A 的异同，其中第8项重申了1.9节定理11和12(a).

231

表 4-1 $m\times n$ 矩阵 A，Col A 和 Nul A 之间的异同

Nul $\boldsymbol{A}$	Col $\boldsymbol{A}$
1. Nul $\boldsymbol{A}$ 是 $\mathbf{R}^n$ 的子空间.	1. Col $\boldsymbol{A}$ 是 $\mathbf{R}^m$ 的子空间.
2. Nul $\boldsymbol{A}$ 是隐式定义的，即你只知道 Nul $\boldsymbol{A}$ 中的向量必须满足某个条件($\boldsymbol{Ax}=\boldsymbol{0}$).	2. Col $\boldsymbol{A}$ 是显式定义的，即你被明确告知如何构造 Col $\boldsymbol{A}$ 中的向量.
3. Nul $\boldsymbol{A}$ 中的向量需经过计算获得，$[\boldsymbol{A}\quad \boldsymbol{0}]$上的行变换是必须的.	3. 找出 Col $\boldsymbol{A}$ 中的向量是非常容易的. $\boldsymbol{A}$ 中的列向量已列出，其他向量可由它们组合而成.
4. 在 Nul $\boldsymbol{A}$ 和 $\boldsymbol{A}$ 的元素之间不存在明显的联系.	4. 在 Col $\boldsymbol{A}$ 和 $\boldsymbol{A}$ 的元素之间的存在明显的联系，因为 $\boldsymbol{A}$ 中的列都在 Col $\boldsymbol{A}$ 中.
5. Nul $\boldsymbol{A}$ 中的向量 $\boldsymbol{v}$ 使得 $\boldsymbol{Av}=\boldsymbol{0}$.	5. Col $\boldsymbol{A}$ 中的向量 $\boldsymbol{v}$ 使得 $\boldsymbol{Ax}=\boldsymbol{v}$ 是相容的.
6. 给定一个具体的向量 $\boldsymbol{v}$，判断 $\boldsymbol{v}$ 是否在 Nul $\boldsymbol{A}$ 中是很容易的，只需计算 $\boldsymbol{Av}$.	6. 给定一个具体的向量 $\boldsymbol{v}$，可以通过计算判断 $\boldsymbol{v}$ 是否在 Col $\boldsymbol{A}$ 中. $[\boldsymbol{A}\quad \boldsymbol{v}]$上的行变换是必须的.
7. Nul $\boldsymbol{A}=\{\boldsymbol{0}\}$当且仅当方程 $\boldsymbol{Ax}=\boldsymbol{0}$ 只有平凡解.	7. Col $\boldsymbol{A}=\mathbf{R}^m$ 当且仅当对 $\mathbf{R}^m$ 中的每个 $\boldsymbol{b}$，方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 都有解.
8. Nul $\boldsymbol{A}=\{\boldsymbol{0}\}$当且仅当线性变换 $\boldsymbol{x}\mapsto\boldsymbol{Ax}$ 是一对一的.	8. Col $\boldsymbol{A}=\mathbf{R}^m$ 当且仅当线性变换 $\boldsymbol{x}\mapsto\boldsymbol{Ax}$ 将 $\mathbf{R}^n$ 映到 $\mathbf{R}^m$ 上.

4.2.5 线性变换的核和值域

除 $\mathbf{R}^n$ 以外，向量空间的子空间通常都用线性变换而非矩阵来描述. 为了使其更精确，我们如下推广 1.8 中给出的定义.

【定义】 从向量空间 V 到 W 内的**线性变换**(linear transformation) T 是满足如下条件的一个从 V 到 W 的对应规则：对 V 中的每个向量 $\boldsymbol{x}$，在 W 中都有唯一的向量 $T(\boldsymbol{x})$ 与之对应，使得

(i) $T(\boldsymbol{u}+\boldsymbol{v})=T(\boldsymbol{u})+T(\boldsymbol{v})$ 对任意 $\boldsymbol{u},\boldsymbol{v}\in V$，且

(ii) $T(c\boldsymbol{u})=cT(\boldsymbol{u})$ 对任意 $\boldsymbol{u}\in V$ 和任意数量 c.

T 的核(kernel)或**零空间**(null space)是 V 中所有满足 $T(\boldsymbol{u})=\boldsymbol{0}$($W$ 中的零向量)的 $\boldsymbol{u}$ 构成的集合. T 的**值域**(range)是 W 中所有形如 $T(\boldsymbol{x})$ 的向量构成的集合，其中 $\boldsymbol{x}\in V$. 如果 T 恰好是以矩阵变换形式出现，即存在某个矩阵 $\boldsymbol{A}$ 使 $T(\boldsymbol{x})=\boldsymbol{Ax}$，则 T 的核和值域恰好是 $\boldsymbol{A}$ 的零空间和列空间，如前面所定义.

232

不难证明 T 的核是 V 的子空间，证明的本质与定理 2 相同. 而且，T 的值域也是 W 的子空间，见图 4-12 和习题 30.

在实际应用中，子空间通常会作为线性变换的核或值域出现. 例如，可以证明一个齐次线性微分方程的解集是一个线性变换的核. 通常，这样的一个线性变换是用函数的一个或多个导数来描述的. 想要详细地说明这一点可能会偏离主题. 所以我们只考虑两个例子，第一个例子解释了微分运算为什么是一个线性变换.

图 4-12 对应于线性变换的子空间

【例题 8】 (需要微积分知识)设 V 是定义在区间$[a,b]$上的所有实值函数构成的向量空间，而且 V 中的元素是可微的并且它们的导数是$[a,b]$上的连续函数. 设 W 是$[a,b]$上全体连续函数构成的向量空间，变换

D: $V \to W$ 将 f 映射到其导数 f'. 微积分中有两条简单的微分法则:

$$D(f+g) = D(f) + D(g) \quad \text{以及} \quad D(cf) = cD(f)$$

它们说明 D 是一个线性变换. 可以证明, D 的核是 $[a, b]$ 上全体常数函数的集合, D 的值域是 $[a, b]$ 上全体连续函数的集合 W. ■

【例题 9】 (需要微积分知识)微分方程

$$y'' + \omega^2 y = 0 \tag{4}$$

公式(4)中的 ω 是一个常数, 它常常用来描述物理系统的变化, 例如弹簧的振动、钟摆的运动以及电磁感应中线圈的电压. (4)的解集是某个线性变换的核, 该变换将函数 $y = f(t)$ 映到函数 $f''(t) + \omega^2 f(t)$. 求解该向量空间的显式表示是一个微分方程问题. 可以证明这个解集是 4.1 节中习题 19 描述的空间. ■

233

基础练习

1. 设 $W = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} : a - 3b - c = 0 \right\}$. 用两种不同的方法证明 W 是 $\mathbf{R}^3$ 的子空间(利用两个定理).

2. 设 $\boldsymbol{A} = \begin{bmatrix} 7 & -3 & 5 \\ -4 & 1 & -5 \\ -5 & 2 & -4 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \boldsymbol{w} = \begin{bmatrix} 7 \\ 6 \\ -3 \end{bmatrix}$. 假设已知 $\boldsymbol{Ax} = \boldsymbol{v}$ 和 $\boldsymbol{Ax} = \boldsymbol{w}$ 都是相容的, 那么关于 $\boldsymbol{Ax} = \boldsymbol{v} + \boldsymbol{w}$, 你能得出何种结论?

习题　4.2

1. 判断是否有 $\boldsymbol{w} = \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix} \in \text{Nul}\,\boldsymbol{A}$, 其中

$$\boldsymbol{A} = \begin{bmatrix} 3 & -5 & -3 \\ 6 & -2 & 0 \\ -8 & 4 & 1 \end{bmatrix}.$$

2. 判断是否有 $\boldsymbol{w} = \begin{bmatrix} 5 \\ -3 \\ 2 \end{bmatrix} \in \text{Nul}\,\boldsymbol{A}$, 其中

$$\boldsymbol{A} = \begin{bmatrix} 5 & 21 & 19 \\ 13 & 23 & 2 \\ 8 & 14 & 1 \end{bmatrix}.$$

在习题 3 ~ 6 中, 列出能张成零空间的向量, 以给出 Nul $\boldsymbol{A}$ 的显式表示.

3. $\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 5 & 0 \\ 0 & 1 & 4 & -2 \end{bmatrix}$

4. $\boldsymbol{A} = \begin{bmatrix} 1 & -6 & 4 & 0 \\ 0 & 0 & 2 & 0 \end{bmatrix}$

5. $\boldsymbol{A} = \begin{bmatrix} 1 & -2 & 0 & 4 & 0 \\ 0 & 0 & 1 & -9 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

6. $\boldsymbol{A} = \begin{bmatrix} 1 & 5 & -4 & -3 & 1 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

在习题 7 ~ 14 中，或者利用恰当的定理证明给定的集合 $\boldsymbol{W}$ 是向量空间，或者找出一个具体的反例.

7. $\left\{\begin{bmatrix} a \\ b \\ c \end{bmatrix} : a+b+c=2\right\}$

8. $\left\{\begin{bmatrix} r \\ s \\ t \end{bmatrix} : 5r-1=s+2t\right\}$

9. $\left\{\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} : \begin{matrix} a-2b=4c \\ 2a=c+3d \end{matrix}\right\}$

10. $\left\{\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} : \begin{matrix} a+3b=c \\ b+c+a=d \end{matrix}\right\}$

11. $\left\{\begin{bmatrix} b-2d \\ 5+d \\ b+3d \\ d \end{bmatrix} : b,d \text{ 是实数}\right\}$

12. $\left\{\begin{bmatrix} b-5d \\ 2b \\ 2d+1 \\ d \end{bmatrix} : b,d \text{ 是实数}\right\}$

13. $\left\{\begin{bmatrix} c-6d \\ d \\ c \end{bmatrix} : c,d \text{ 是实数}\right\}$

14. $\left\{\begin{bmatrix} -a+2b \\ a-2b \\ 3a-6b \end{bmatrix} : a,b \text{ 是实数}\right\}$

在习题 15 和 16 中，求 $\boldsymbol{A}$ 使得给定的集合是 Col $\boldsymbol{A}$.

15. $\left\{\begin{bmatrix} 2s+3t \\ r+s-2t \\ 4r+s \\ 3r-s-t \end{bmatrix} : r,s,t \text{ 是实数}\right\}$

16. $\left\{\begin{bmatrix} b-c \\ 2b+c+d \\ 5c-4d \\ d \end{bmatrix} : b,c,d \text{ 是实数}\right\}$

对习题 17 ~ 20 中的向量，(a)求出 k，使得 Nul $\boldsymbol{A}$ 是 $\mathbf{R}^k$ 的一个子空间，(b)求出 k，使得 Col $\boldsymbol{A}$ 是 $\mathbf{R}^k$ 的一个子空间. [234]

17. $\boldsymbol{A}=\begin{bmatrix} 2 & -6 \\ -1 & 3 \\ -4 & 12 \\ 3 & -9 \end{bmatrix}$

18. $\boldsymbol{A}=\begin{bmatrix} 7 & -2 & 0 \\ -2 & 0 & -5 \\ 0 & -5 & 7 \\ -5 & 7 & -2 \end{bmatrix}$

19. $\boldsymbol{A}=\begin{bmatrix} 4 & 5 & -2 & 6 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}$

20. $\boldsymbol{A}=\begin{bmatrix} 1 & -3 & 9 & 0 & -5 \end{bmatrix}$

21. 设 $\boldsymbol{A}$ 与习题 17 中相同，试分别找出 Nul $\boldsymbol{A}$ 和 Col $\boldsymbol{A}$ 中的一个非零向量.

22. 设 $\boldsymbol{A}$ 与习题 3 中相同，试分别找出 Nul $\boldsymbol{A}$ 和 Col $\boldsymbol{A}$ 中的一个非零向量.

23. 设 $\boldsymbol{A}=\begin{bmatrix} -6 & 12 \\ -3 & 6 \end{bmatrix}, \boldsymbol{w}=\begin{bmatrix} 2 \\ 1 \end{bmatrix}$. 判断 $\boldsymbol{w}$ 是否属于 Col $\boldsymbol{A}$ 和 Nul $\boldsymbol{A}$?

24. 设 $\boldsymbol{A}=\begin{bmatrix} -8 & -2 & -9 \\ 6 & 4 & 8 \\ 4 & 0 & 4 \end{bmatrix}, \boldsymbol{w}=\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$. 判断 $\boldsymbol{w}$ 是否属于 Col $\boldsymbol{A}$ 和 Nul $\boldsymbol{A}$?

在习题 25 和 26 中，判断各个命题的真假，并说明理由，其中 $\boldsymbol{A}$ 表示 $m\times n$ 矩阵.

25. a. $\boldsymbol{A}$ 的零空间是方程 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解集.
    b. $m\times n$ 矩阵的零空间属于 $\mathbf{R}^m$.
    c. $\boldsymbol{A}$ 的列空间是映射 $\boldsymbol{x}\to\boldsymbol{Ax}$ 的值域.
    d. 如果 $\boldsymbol{Ax}=\boldsymbol{b}$ 是相容的，则 Col $\boldsymbol{A}$ 就是 $\mathbf{R}^m$.

e. 线性变换的核是一个向量空间.

f. Col $\boldsymbol{A}$ 是所有可以写成 $\boldsymbol{Ax}$ 的向量构成的集合.

26. a. 零空间是向量空间.

b. $m\times n$ 矩阵的列空间属于 $\mathbf{R}^m$.

c. Col $\boldsymbol{A}$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的所有解组成的集合.

d. Nul $\boldsymbol{A}$ 是映射 $\boldsymbol{x}\to\boldsymbol{Ax}$ 的核.

e. 线性变换的值域是向量空间.

f. 齐次线性微分方程的解集合是一个线性变换的核.

27. 可以证明下面方程组的一个解是 $x_2=3, x_2=2, x_3=-1$. 利用该事实和这节的理论来解释为何另一个解是 $x_1=30, x_2=20, x_3=-10$（不作计算,观察解之间的关系）.

$$\begin{aligned} x_1-3x_2-3x_3&=0\\ -2x_1+4x_2+2x_3&=0\\ -x_1+5x_2+7x_3&=0 \end{aligned}$$

28. 考虑下列两个方程组：

$$\begin{aligned} 5x_1+x_2-3x_3&=0\\ -9x_1+2x_2+5x_3&=1\\ 4x_1+x_2-6x_3&=9 \end{aligned}\qquad\qquad \begin{aligned} 5x_1+x_2-3x_3&=0\\ -9x_1+2x_2+5x_3&=5\\ 4x_1+x_2-6x_3&=45 \end{aligned}$$

可以证明前一个方程组有解. 利用该事实和本节的理论来解释为什么后一个方程组也必有一个解（不用行变换）.

29. 定理3可按下列步骤证明：给定一个 $m\times n$ 矩阵 $\boldsymbol{A}$，则 Col $\boldsymbol{A}$ 中的元素可以写成 $\boldsymbol{Ax}$，其中 $\boldsymbol{x}\in\mathbf{R}^n$. 设 $\boldsymbol{Ax}$ 和 $\boldsymbol{Aw}$ 代表 Col $\boldsymbol{A}$ 中的任意两个向量.

a. 解释零向量为什么属于 Col $\boldsymbol{A}$.

b. 证明：$\boldsymbol{Ax}+\boldsymbol{Aw}\in$ Col $\boldsymbol{A}$.

c. 给定一个数量 c，证明：$c(\boldsymbol{Ax})\in$ Col $\boldsymbol{A}$.

30. 设 T: $V\to W$ 是从向量空间 V 到 W 的线性变换. 证明：T 的值域是 W 的一个子空间. [提示：值域中的典型元素都有 $\boldsymbol{Ax}$ 形式，对于某个 V 中的 $\boldsymbol{x}$.]

31. 定义 $T:\mathbf{P}_2\to\mathbf{R}^2$ 为 $T(\boldsymbol{p})=\begin{bmatrix}\boldsymbol{p}(0)\\ \boldsymbol{p}(1)\end{bmatrix}$. 例如，如果 $\boldsymbol{p}(t)=3+5t+7t^2$，则 $T(\boldsymbol{p})=\begin{bmatrix}3\\15\end{bmatrix}$.

a. 证明：T 是一个线性变换. [提示：对任意 $\mathbf{P}_2$ 中的多项式 $\boldsymbol{p}$，$\boldsymbol{q}$，计算 $T(\boldsymbol{p}+\boldsymbol{q})$ 和 $T(c\boldsymbol{p})$.]

b. 试求 $\mathbf{P}_2$ 中的一多项式 $\boldsymbol{p}$，使其可以张成 T 的核，并且刻画 T 的值域.

32. 定义线性变换 T：$\mathbf{P}_2\to\mathbf{R}^2$ 为 $T(p)=\begin{bmatrix}\boldsymbol{p}(0)\\ \boldsymbol{p}(0)\end{bmatrix}$. 试求 $\mathbf{P}_2$ 中的多项式 $\boldsymbol{p}_1$，$\boldsymbol{p}_2$，使它们张成 T 的核，并且刻画 T 的值域.

33. 设 $M_{2\times2}$ 是所有 2×2 矩阵构成的向量空间，并且定义 $T:M_{2\times2}\to M_{2\times2}$ 为：$T(\boldsymbol{A})=\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}$，其中 $\boldsymbol{A}=\begin{bmatrix}a&b\\c&d\end{bmatrix}$.

a. 证明：T 是一个线性变换.

235

b. 设 $\boldsymbol{B}$ 是 $M_{2\times2}$ 中满足 $\boldsymbol{B}^{\mathrm{T}}=\boldsymbol{B}$ 的元素，试在 $M_{2\times2}$ 中求满足 $T(\boldsymbol{A})=\boldsymbol{B}$ 的一个 $\boldsymbol{A}$.

c. 证明：T 的值域是 $M_{2\times2}$ 中所有满足性质 $\boldsymbol{B}^{\mathrm{T}}=\boldsymbol{B}$ 的 $\boldsymbol{B}$ 构成的集合.

d. 描述 T 的核.

34. （需要微积分知识）如下定义 T：$C[0, 1]\to C[0, 1]$：对 $C[0, 1]$ 中的 f，设 $T(f)$ 是 f 的不

定积分 F，并且满足 $F(0)=0$. 证明：T 是一个线性变换，并且描述 T 的核(见4.1节习题20中的记号).

35. 设 V 和 W 是向量空间，T: $V\to W$ 是线性变换. 给定 V 的一个子空间 U，设 $T(U)$ 表示全体像 $T(\boldsymbol{x})$ 组成的集合，其中 $\boldsymbol{x}\in U$. 证明：$T(U)$ 是 W 的子空间.
36. 给定 T: $V\to W$ 同习题35，Z 是 W 的一个子空间. 设 U 是 V 中所有使得 $T(\boldsymbol{x})\in Z$ 的 $\boldsymbol{x}$ 组成的集合. 证明：U 是 W 的子空间.
37. [**M**]判断 $\boldsymbol{w}$ 是否属于 $\boldsymbol{A}$ 的列空间或者 $\boldsymbol{A}$ 的零空间，或者同时属于这两个空间，其中

$$\boldsymbol{w}=\begin{bmatrix}1\\1\\-1\\-3\end{bmatrix},\quad A=\begin{bmatrix}7&6&-4&1\\-5&-1&0&-2\\9&-11&7&-3\\19&-9&7&1\end{bmatrix}$$

38. [**M**]判断 $\boldsymbol{w}$ 是否属于 $\boldsymbol{A}$ 的列空间或者 $\boldsymbol{A}$ 的零空间，或者同时属于这两个空间，其中

$$\boldsymbol{w}=\begin{bmatrix}1\\2\\1\\0\end{bmatrix},\quad A=\begin{bmatrix}-8&5&-2&0\\-5&2&1&-2\\10&-8&6&-3\\3&-2&1&0\end{bmatrix}$$

39. [**M**]以 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_5$ 记矩阵 $\boldsymbol{A}$ 的列向量，其中

$$\boldsymbol{A}=\begin{bmatrix}5&1&2&2&0\\3&3&2&-1&-12\\8&4&4&-5&12\\2&1&1&0&-2\end{bmatrix},\quad \boldsymbol{B}=[\boldsymbol{a}_1\quad \boldsymbol{a}_2\quad \boldsymbol{a}_4]$$

a. 解释 $\boldsymbol{a}_3$ 和 $\boldsymbol{a}_5$ 为什么属于 $\boldsymbol{B}$ 的列空间.

b. 求可以张成 Nul $\boldsymbol{A}$ 的向量集合.

c. 设 T: $\mathbf{R}^5\to\mathbf{R}^4$ 定义为 $T(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{x}$. 解释 T 为什么既不是一对一的，也不是到上的.

40. [**M**]设 $H=\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$, $K=\text{Span}\{\boldsymbol{v}_3,\boldsymbol{v}_4\}$，其中

$$\boldsymbol{v}_1=\begin{bmatrix}5\\3\\8\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}1\\3\\4\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}2\\-1\\5\end{bmatrix},\boldsymbol{v}_4=\begin{bmatrix}0\\-12\\-28\end{bmatrix}$$

则 H 和 K 是 $\mathbf{R}^3$ 的子空间. 实际上，H 和 K 是 $\mathbf{R}^3$ 中过原点的平面，它们的交是过 $\mathbf{0}$ 的直线. 试求生成该直线的非零向量 $\boldsymbol{w}$. [*提示*：$\boldsymbol{w}$ 可以写成 $c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2$，也可写成 $c_3\boldsymbol{v}_3+c_4\boldsymbol{v}_4$. 为构造 $\boldsymbol{w}$，只需解方程 $c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2=c_3\boldsymbol{v}_3+c_4\boldsymbol{v}_4$ 即可.]

基础练习答案

1. *第一种方法*：因为 W 是线性齐次方程组(这里的方程组只含有一个方程)的全部解的集合，由定理2知 W 是 $\mathbf{R}^3$ 的一个子空间. 等价地，W 是 1×3 矩阵 $\boldsymbol{A}=[1\quad -3\quad -1]$ 的零空间.

第二种方法：对首变量 a，根据自由变量 b 和 c，解方程 $a-3b-c=0$. 任意解都有 $\begin{bmatrix}3b+c\\b\\c\end{bmatrix}$ 的形式，其中 b 和 c 任意，并且

$$\begin{bmatrix}3b+c\\b\\c\end{bmatrix}=b\begin{bmatrix}3\\1\\0\end{bmatrix}+c\begin{bmatrix}1\\0\\1\end{bmatrix}$$

$$\qquad\quad\uparrow\qquad\ \ \uparrow$$
$$\qquad\quad\boldsymbol{v}_1\qquad\ \boldsymbol{v}_2$$

这表明 $W = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$. 因此由定理1知 $\boldsymbol{W}$ 是 $\mathbf{R}^3$ 的一个子空间. 我们也可以对 b 或 c 解方程 $a-3b-c=0$, 并且将 W 表示成另两个向量的线性组合构成的集合.

2. $\boldsymbol{v}$ 和 $\boldsymbol{w}$ 都在 Col $\boldsymbol{A}$ 中, 因为 Col $\boldsymbol{A}$ 是向量空间, $\boldsymbol{v}+\boldsymbol{w}$ 一定在 Col $\boldsymbol{A}$ 中, 即方程 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{v}+\boldsymbol{w}$ 是相容的.

4.3 线性无关集: 基

本节中, 我们将认识并且研究能够"最高效地"张成向量空间 V 或者子空间 H 的子集, 其关键思想是线性无关性, 它是 $\mathbf{R}^n$ 中线性无关性质的推广.

称 V 中带指标的向量集 $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p\}$ **线性无关**(linearly independent), 如果向量方程

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_p\boldsymbol{v}_p = \boldsymbol{0} \tag{1}$$

只有平凡解 $c_1 = 0, \cdots, c_p = 0$.[1]

称集合 $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p\}$ **线性相关**(linearly dependent), 如果(1)有非平凡解, 即存在 $c_1, \cdots, c_p$ 不全为零, 使(1)成立. 在这种情况下, 称(1)是 $v_1, \cdots, v_p$ 之间的**线性相关关系**(linear dependence relation).

正如在 $\mathbf{R}^n$ 中, 只含一个向量 $\boldsymbol{v}$ 的集合是线性无关的当且仅当 $\boldsymbol{v} \neq \boldsymbol{0}$. 同样, 含两个向量的集合是线性相关的当且仅当一个是另一个的倍数. 任意包含零向量的集合是线性相关的. 下面定理的证明与1.7节定理7的证明相同.

【定理4】 设加标集 $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p\}$ 至少含有两个向量, 并且 $\boldsymbol{v}_1 \neq \boldsymbol{0}$, 则 $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p\}$ 是线性相关的当且仅当某个 $\boldsymbol{v}_j (j>1)$ 是其前面向量 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{j-1}$ 的线性组合.

$\mathbf{R}^n$ 空间与一般向量空间的线性相关的主要不同在于, 当向量不是 n 元组时, 齐次方程(1)通常不能写成含 n 个线性方程的方程组, 即不可能将这些向量写成某个矩阵 $\boldsymbol{A}$ 的列, 然后转而研究 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$. 我们必须依赖线性相关的定义和定理4.

【例题1】 设 $\boldsymbol{p}_1(t)=1, \boldsymbol{p}_2(t)=t$ 且 $\boldsymbol{p}_3(t)=4-t$, 则 $\{\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3\}$ 是线性相关的, 因为 $\boldsymbol{p}_3 = 4\boldsymbol{p}_1 - \boldsymbol{p}_2$. ■

【例题2】 在 $C[0,1]$(所有 $0 \leqslant t \leqslant 1$ 上的连续函数构成的向量空间)中, 集合 $\{\sin t, \cos t\}$ 是线性无关的. 因为作为 $C[0,1]$ 中的向量, $\sin t$ 和 $\cos t$ 相互之间无倍数关系, 即, 不存在数量 c 使得对 $[0, 1]$ 中的任意 t 都有 $\cos t = c \cdot \sin t$. (观察 $\sin t$ 与 $\cos t$ 的图像.)然而, $\{\sin t \cos t, \sin 2t\}$ 是线性相关的, 因为对任意 t 有恒等式: $\sin 2t = 2\sin t \cos t$. ■

237

【定义】 设 H 是向量空间 V 的子空间, 称 V 中向量的加标集 $\mathcal{B} = \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_p\}$ 是 H 的一组**基**(basis), 如果

(i) $\mathcal{B}$ 是一个线性无关集, 并且

(ii) 由 $\mathcal{B}$ 张成的子空间与 H 相同, 即

$$H = \text{Span}\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_p\}$$

当 $H=V$ 时, 基的定义同样适用, 因为任何向量空间都是其自身的子空间. 因此

1. 与第1章一样, 为方便起见, 在(1)中我们用 $c_1, \cdots, c_p$ 代替 $x_1, \cdots, x_p$.

V的基是一个能够张成V的线性无关集. 当$H \neq V$时，(ii)包含必要条件：$\boldsymbol{b}_1, \cdots, \boldsymbol{b}_p$中的每个向量一定属于$H$，因为在4.1节我们已经知道Span$\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_p\}$包含$\boldsymbol{b}_1, \cdots, \boldsymbol{b}_p$.

【例题3】 设$\boldsymbol{A}$是可逆的$n \times n$矩阵，即$\boldsymbol{A} = [\boldsymbol{a}_1 \cdots \boldsymbol{a}_n]$. 那么$\boldsymbol{A}$的列向量组成$\mathbf{R}^n$的一组基，这时由可逆矩阵的定理知，它们线性无关并且张成$\mathbf{R}^n$. ■

【例题4】 设$\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$是$n \times n$单位矩阵$\boldsymbol{I}_n$的列向量，即

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \cdots, \boldsymbol{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

称集合$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$是$\mathbf{R}^n$的**标准基**(standard basis)(见图4-13).

【例题5】 设$\boldsymbol{v}_1 = \begin{bmatrix} 3 \\ 0 \\ -6 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 7 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix}$. 判断$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$是否是$\mathbf{R}^n$的基.

解： 因为这里恰好有三个$\mathbf{R}^3$中的向量，我们可用多种方法判断矩阵$\boldsymbol{A} = [\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \boldsymbol{v}_3]$是否可逆. 例如，经过两个行替换可知$\boldsymbol{A}$有三个主元位置. 因此$\boldsymbol{A}$是可逆的. 与例题3一样，$\boldsymbol{A}$的列向量构成$\mathbf{R}^3$的一组基. ■

【例题6】 设$S = \{1, t, t^2, \cdots, t^n\}$. 验证$S$是$\mathbf{P}_n$的一组基. 称这组基为$\mathbf{P}_n$的**标准基**(standard basis). ■ 238

解： 显然S张成$\mathbf{P}_n$. 为证明S是线性无关的，假设$c_0, \cdots, c_n$满足

$$c_0 \cdot 1 + c_1 t + c_2 t^2 + \cdots + c_n t^n = \mathbf{0}(t) \tag{2}$$

该等式意味着左端多项式与右端零多项式取值相同. 代数基本定理说明$\mathbf{P}_n$中含有多于n个零点的多项式只能是零多项式. 即(2)对任意t成立当且仅当$c_0 = \cdots = c_n = 0$. 这就证明了S是线性无关的，因此是$\mathbf{P}_n$的一组基. 见图4-14. ■

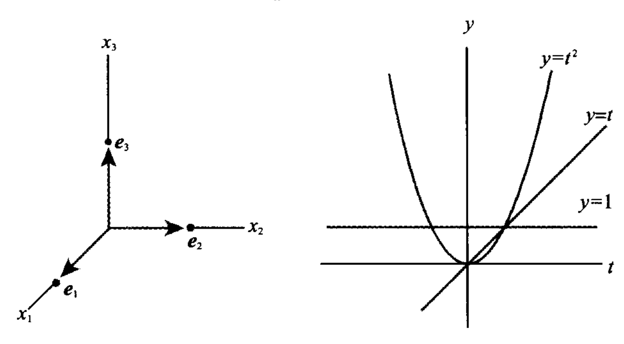


图4-13 $\mathbf{R}^n$的标准基　　图4-14 $\mathbf{P}_2$的标准基

在4.4节会介绍一种方法，它能够更好地处理$\mathbf{P}_n$中关于线性无关和扩张的问题.

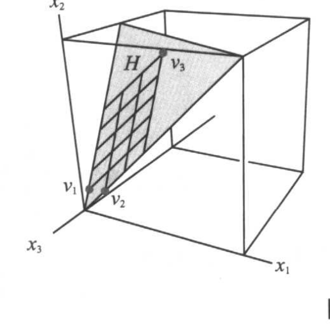
4.3.1 张成集定理

我们即将看到，基是一组“有效”的张成集，其中不含非必要的向量. 实际上，可以通过去掉一个张成集中的非必要的向量而得到基.

【例题7】　设 $\boldsymbol{v}_1=\begin{bmatrix}0\\2\\-1\end{bmatrix}, \boldsymbol{v}_2=\begin{bmatrix}2\\2\\0\end{bmatrix}, \boldsymbol{v}_3=\begin{bmatrix}6\\16\\-5\end{bmatrix}$，并且 $H=\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$. 注意 $\boldsymbol{v}_3=5\boldsymbol{v}_1+3\boldsymbol{v}_2$，证明 $\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}=\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$，然后求子空间 H 的一组基.

解： $\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$中的每个向量属于 H，因为

$$c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+0\boldsymbol{v}_3$$



现在，设 $\boldsymbol{x}$ 是 H 中的任意向量，假定 $\boldsymbol{x}=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+c_3\boldsymbol{v}_3$. 因为 $\boldsymbol{v}_3=5\boldsymbol{v}_1+3\boldsymbol{v}_2$，我们可以做如下替换

$$\begin{aligned}\boldsymbol{x}&=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+c_3(5\boldsymbol{v}_1+3\boldsymbol{v}_2)\\&=(c_1+5c_3)\boldsymbol{v}_1+(c_2+3c_3)\boldsymbol{v}_2\end{aligned}$$

因此 $\boldsymbol{x}\in\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$. 故 H 中的每个向量都属于 $\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$. 我们可以断定 H 和$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$是同一个向量集. 又因为$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$显然是线性无关的，可知$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$是 H 的基. ■

下面的定理是例题7的推广.

【定理5】　张成集定理

设 $S=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$是 V 中的一个集合，$H=\text{Span}\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$.

(a) 如果 S 中的一个向量，比如说 $\boldsymbol{v}_k$，是 S 中其他向量的线性组合，则将 $\boldsymbol{v}_k$从 S 中删除后得到的集合仍然可以张成 H.

(b) 如果 $H\neq\{\boldsymbol{0}\}$，则 S 的某个子集是 H 的基.

239

证明：

(a)通过重新排列 S 中向量的次序，我们可以假设 $\boldsymbol{v}_p$是 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{p-1}$的线性组合，也就是说

$$\boldsymbol{v}_p=a_1\boldsymbol{v}_1+\cdots+a_{p-1}\boldsymbol{v}_{p-1} \tag{3}$$

给定 H 中的任意一个 $\boldsymbol{x}$，选取适当的 $c_1,\cdots,c_p$，我们有

$$\boldsymbol{x}=c_1\boldsymbol{v}_1+\cdots+c_{p-1}\boldsymbol{v}_{p-1}+c_p\boldsymbol{v}_p \tag{4}$$

用(3)的表示替换(4)中的 $\boldsymbol{v}_p$，可知 $\boldsymbol{x}$ 是 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{p-1}$的线性组合. 由于 $\boldsymbol{x}$ 是 H 中的任意一个元素，所以$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{p-1}\}$张成 H.

(b)如果最初的张成集 S 是线性无关的，它已经是 H 的一组基. 否则，如果 S 中的一个向量依赖于其他的向量，则由(a)该向量可以删除掉. 只要张成集中不少于两个向量，就可以重复上述过程，直到变成一个线性无关集，进而是 H 的基. 如果张成集最终减少到一个向量，那么该向量必是非零向量(因此是线性无关的)，这是由于 $H\neq\{\boldsymbol{0}\}$. ■

4.3.2　Nul A 和 Col A 的基

我们已经知道如何求出张成矩阵 $\boldsymbol{A}$ 的零空间的向量. 4.2节中的讨论指出，我们的方法已经可以产生一个线性无关集. 因此这种方法可以产生 Nul $\boldsymbol{A}$ 的基.

下面的两个例题描述了一种求列空间基的简单算法.

【例题8】　求 Col $\boldsymbol{B}$ 的一组基，其中

$$B=[b_1 \quad b_2 \quad \cdots \quad b_5]=\begin{bmatrix}1&4&0&2&0\\0&0&1&-1&0\\0&0&0&0&1\\0&0&0&0&0\end{bmatrix}$$

解：B 的每个非主元列都是主元列的线性组合. 实际上，$b_2=4b_1, b_4=2b_1-b_3$. 由张成集定理，我们可以去掉 b_2 和 b_4，$\{b_1,b_3,b_5\}$仍可以张成 Col B. 令

$$S=\{b_1,b_3,b_5\}=\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\\0\end{bmatrix}\right\}$$

由于 $b_1\neq 0$，并且 S 中没有向量是其他向量的线性组合，S 是线性无关的(定理 4). 因此 S 是 Col B 的一组基. ■ 240

如果 A 不是简化阶梯矩阵该怎么办? 回顾前面的知识，A 的列向量之间的任何线性相关关系都可以表示成 $Ax=0$ 的形式，其中 x 是一列权重(如果某些列不包含在某个线性相关关系当中,则它们的权重为零). 当 A 通过行化简转化成矩阵 B 时，B 的列常常与 A 的列截然不同. 但是，$Ax=0$ 和 $Bx=0$ 有完全相同的解集，换句话说，A 的列与 B 的列具有相同的线性相关关系.

矩阵的初等行变换不会影响列向量的线性相关关系.

【例题 9】 可以证明

$$A=[a_1 \quad a_2 \quad \cdots \quad a_5]=\begin{bmatrix}1&4&0&2&-1\\3&12&1&5&5\\2&8&1&3&2\\5&20&2&8&8\end{bmatrix}$$

与例题 8 中的矩阵 B 行等价. 试求 Col A 的一组基.

解：在例题 8 中我们看到

$$b_2=4b_1 \text{且} b_4=2b_1-b_3$$

所以我们期望

$$a_2=4a_1 \text{且} a_4=2a_1-a_3$$

经检验事实确实如此! 因此选取 Col A 的极小张成集时，可以去掉 a_2和 a_4. 实际上，$\{a_1,a_3,a_5\}$一定是线性无关的，这是因为 a_1,a_3,a_5之间的任何线性相关关系都会导致 b_1,b_3,b_5之间的一个线性相关关系. 但我们又知道$\{b_1,b_3,b_5\}$是线性无关的，因此$\{a_1,a_3,a_5\}$是 Col A 的一组基. 这个基中的列向量都是 A 中的主元列. ■

例题 8 和 9 说明了下面这个有用的事实.

【定理 6】 矩阵 A 中的主元列组成 Col A 的一组基.

证明：一般的证明要用到上面的讨论. 设 B 是 A 的简化阶梯形式. B 的主元列所构成的集合是线性无关的，即，集合中的任意向量都不是前面向量的线性组合. 因为 A 行等价于 B，A 中列的任意线性相关关系都与 B 中列的一个线性相关关系相对应，所

以 $\boldsymbol{A}$ 的主元列也是线性无关的. 同理, $\boldsymbol{A}$ 的每个非主元列是 $\boldsymbol{A}$ 中主元列的线性组合, 241 因此, 由张成集定理知 $\boldsymbol{A}$ 的非主元列就可以从 Col $\boldsymbol{A}$ 的张成集中去掉, Col $\boldsymbol{A}$ 的基中就只剩下 $\boldsymbol{A}$ 的主元列. ■

警告: 需要注意, 构造 Col $\boldsymbol{A}$ 的基时应当使用 $\boldsymbol{A}$ 本身的主元列. 行变换会改变矩阵的列空间. $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的简化阶梯形, 但 $\boldsymbol{B}$ 中的列向量通常不在 $\boldsymbol{A}$ 的列空间中. 例如, 例题 8 中 $\boldsymbol{B}$ 的列向量, 它们最后一个分量都是零, 因此不可能张成例题 9 中矩阵 $\boldsymbol{A}$ 的列空间.

对基的两种理解

运用张成集定理时, 一旦张成集线性无关, 我们就应该立刻停止删除向量. 如果删掉了一个不能写成其他向量线性组合的向量, 则这个小的集合就不能张成 V. 因此基是一个极小张成集.

基同时也是一个极大的线性无关集. 设 S 是 V 的一组基, 如果 S 中再增加一个 V 中的向量, 比如说 $\boldsymbol{w}$, 则新集合就不再线性无关, 因为 S 张成 V, $\boldsymbol{w}$ 是 S 中元素的一个线性组合.

【例题 10】　$\mathbf{R}^3$ 中的下列 3 个集合演示了一个线性无关集如何扩张成基, 以及进一步的扩张如何破坏该集合的无关性. 同样, 一个张成集可以收缩为基, 但进一步的收缩也会破坏张成性质.

$$\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix}\right\}\qquad\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix}\right\}\qquad\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix},\begin{bmatrix}7\\8\\9\end{bmatrix}\right\}$$

线性无关但不能张成 $\mathbf{R}^3$　　　$\mathbf{R}^3$ 的一组基　　　可以张成 $\mathbf{R}^3$ 但线性相关

基础练习

1. 设 $\boldsymbol{v}_1=\begin{bmatrix}1\\-2\\3\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}-2\\7\\-9\end{bmatrix}$. 判断 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是否是 $\mathbf{R}^3$ 的一组基. $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是否是 $\mathbf{R}^2$ 的基?

2. 设 $\boldsymbol{v}_1=\begin{bmatrix}1\\-3\\4\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}6\\2\\-1\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}2\\-2\\3\end{bmatrix},\boldsymbol{v}_4=\begin{bmatrix}-4\\-8\\9\end{bmatrix}$. 求由 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4\}$ 张成的子空间 W 的一242组基.

3. 设 $\boldsymbol{v}_1=\begin{bmatrix}1\\0\\0\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}0\\1\\0\end{bmatrix}$, 并且 $H=\left\{\begin{bmatrix}s\\s\\0\end{bmatrix}: s\in\mathbf{R}\right\}$, 则 H 中的每个向量都是 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的线性组合, 这是由于 $\begin{bmatrix}s\\s\\0\end{bmatrix}=s\begin{bmatrix}1\\0\\0\end{bmatrix}+s\begin{bmatrix}0\\1\\0\end{bmatrix}$. $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是否是 H 的基?

习题　4.3

判断习题 1 ~ 8 中哪些集合是 $\mathbf{R}^3$ 的基. 对不是基的集合，判断哪些是线性无关的，哪些可以张成 $\mathbf{R}^3$. 说明理由.

1. $\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}1\\1\\1\end{bmatrix}$

2. $\begin{bmatrix}1\\0\\1\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix}$

3. $\begin{bmatrix}1\\0\\-2\end{bmatrix},\begin{bmatrix}3\\2\\-4\end{bmatrix},\begin{bmatrix}-3\\-5\\1\end{bmatrix}$

4. $\begin{bmatrix}2\\-2\\1\end{bmatrix},\begin{bmatrix}1\\-3\\2\end{bmatrix},\begin{bmatrix}-7\\5\\4\end{bmatrix}$

5. $\begin{bmatrix}1\\-3\\0\end{bmatrix},\begin{bmatrix}-2\\9\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}0\\-3\\5\end{bmatrix}$

6. $\begin{bmatrix}1\\2\\-3\end{bmatrix},\begin{bmatrix}-4\\-5\\6\end{bmatrix}$

7. $\begin{bmatrix}-2\\3\\0\end{bmatrix},\begin{bmatrix}6\\-1\\5\end{bmatrix}$

8. $\begin{bmatrix}1\\-4\\3\end{bmatrix},\begin{bmatrix}0\\3\\-1\end{bmatrix},\begin{bmatrix}3\\-5\\4\end{bmatrix},\begin{bmatrix}0\\2\\-2\end{bmatrix}$

对于习题 9 和 10 给出的矩阵，求它们的零空间的基，参考 4.2 节例题 3 后的讨论.

9. $\begin{bmatrix}1 & 0 & -3 & 2\\0 & 1 & -5 & 4\\3 & -2 & 1 & -2\end{bmatrix}$

10. $\begin{bmatrix}1 & 0 & -5 & 1 & 4\\-2 & 1 & 6 & -2 & -2\\0 & 2 & -8 & 1 & 9\end{bmatrix}$

11. 求 $\mathbf{R}^3$ 中平面 $x+2y+z=0$ 上的向量的基. [提示：把方程看作一个齐次线性方程组.]

12. 求 $\mathbf{R}^2$ 中直线 $y=5x$ 上的向量的基.

在习题 13 和 14 中，$\boldsymbol{A}$ 行等价于 $\boldsymbol{B}$，分别求 Nul $\boldsymbol{A}$ 和 Col $\boldsymbol{A}$ 的基.

13. $\boldsymbol{A}=\begin{bmatrix}-2 & 4 & -2 & -4\\2 & -6 & -3 & 1\\-3 & 8 & 2 & -3\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}1 & 0 & 6 & 5\\0 & 2 & 5 & 3\\0 & 0 & 0 & 0\end{bmatrix}$

14. $\boldsymbol{A}=\begin{bmatrix}1 & 2 & -5 & 11 & -3\\2 & 4 & -5 & 15 & 2\\1 & 2 & 0 & 4 & 5\\3 & 6 & -5 & 19 & -2\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}1 & 2 & 0 & 4 & 5\\0 & 0 & 5 & -7 & 8\\0 & 0 & 0 & 0 & -9\\0 & 0 & 0 & 0 & 0\end{bmatrix}$

在习题 15 ~ 18 中，求由给定的向量张成的空间的基.

15. $\begin{bmatrix}1\\0\\-3\\2\end{bmatrix},\begin{bmatrix}0\\1\\2\\-3\end{bmatrix},\begin{bmatrix}-3\\-4\\1\\6\end{bmatrix},\begin{bmatrix}1\\-3\\-8\\7\end{bmatrix},\begin{bmatrix}2\\1\\-6\\9\end{bmatrix}$

16. $\begin{bmatrix}1\\0\\0\\1\end{bmatrix},\begin{bmatrix}-2\\1\\-1\\1\end{bmatrix},\begin{bmatrix}6\\-1\\2\\-1\end{bmatrix},\begin{bmatrix}5\\-3\\3\\-4\end{bmatrix},\begin{bmatrix}0\\3\\-1\\1\end{bmatrix}$

17. [M] $\begin{bmatrix}8\\9\\-3\\-6\\0\end{bmatrix},\begin{bmatrix}4\\5\\1\\-4\\4\end{bmatrix},\begin{bmatrix}-1\\-4\\-9\\6\\-7\end{bmatrix},\begin{bmatrix}6\\8\\4\\-7\\10\end{bmatrix},\begin{bmatrix}-1\\4\\11\\-8\\-7\end{bmatrix}$

18. [M] $\begin{bmatrix}-8\\7\\6\\5\\-7\end{bmatrix},\begin{bmatrix}8\\-7\\-9\\-5\\7\end{bmatrix},\begin{bmatrix}-8\\7\\4\\5\\-7\end{bmatrix},\begin{bmatrix}1\\4\\9\\6\\-7\end{bmatrix},\begin{bmatrix}-9\\3\\-4\\-1\\0\end{bmatrix}$

243

19. 设 $\boldsymbol{v}_1=\begin{bmatrix}4\\-3\\7\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}1\\9\\-2\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}7\\11\\6\end{bmatrix},H=\mathrm{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$，可以验证 $4\boldsymbol{v}_1+5\boldsymbol{v}_2-3\boldsymbol{v}_3=\mathbf{0}$，利

用这些信息求 H 的一组基．答案不唯一．

20. 设 $\mathbf{v}_1=\begin{bmatrix}7\\4\\-9\\-5\end{bmatrix},\mathbf{v}_2=\begin{bmatrix}4\\-7\\2\\5\end{bmatrix},\mathbf{v}_3=\begin{bmatrix}1\\-5\\3\\4\end{bmatrix}$．可以验证 $\mathbf{v}_1-3\mathbf{v}_2+5\mathbf{v}_3=\mathbf{0}$，利用这些信息求 $H=\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$的一组基．

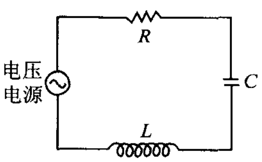
在习题21和22中，判断各命题的真假，并说明理由．

21. a. 单个的向量是线性相关的．
    b. 如果 $H=\mathrm{Span}\{\mathbf{b}_1,\cdots,\mathbf{b}_p\}$，则$\{\mathbf{b}_1,\cdots,\mathbf{b}_p\}$是 H 的一组基．
    c. 一个 $n\times n$ 可逆矩阵的列向量构成 $\mathbf{R}^n$的一组基．
    d. 一组基就是一个极大张成集．
    e. 在有些情况下，矩阵的列之间的线性相关关系会受到某些初等行变换的影响．
22. a. 子空间 H 中的线性无关集是 H 的基．
    b. 如果非零向量的有限集合 S 可以张成向量空间 V，则 S 的某个子集是 V 的一组基．
    c. 基是一个极大线性无关集．
    d. 4.2 节中描述的产生 Nul A 的基的标准方法，有时不能产生 Nul A 的基．
    e. 如果 B 是 A 的阶梯矩阵，则 B 的主元列构成 Col A 的一组基．
23. 假设 $\mathbf{R}^4=\mathrm{Span}\{\mathbf{v}_1,\cdots,\mathbf{v}_4\}$，解释$\{\mathbf{v}_1,\cdots,\mathbf{v}_4\}$为什么是 $\mathbf{R}^4$ 的一组基．
24. 设 $\mathcal{B}=\{\mathbf{v}_1,\cdots,\mathbf{v}_n\}$是 $\mathbf{R}^n$ 中的一个线性无关集，解释 $\mathcal{B}$ 为什么一定是 $\mathbf{R}^n$的一组基．
25. 设 $\mathbf{v}_1=\begin{bmatrix}1\\0\\1\end{bmatrix},\mathbf{v}_2=\begin{bmatrix}0\\1\\1\end{bmatrix},\mathbf{v}_3=\begin{bmatrix}0\\1\\0\end{bmatrix}$，$H$ 是 $\mathbf{R}^3$ 中第二个和第三个分量相同的向量的集合，则 H 中的每个向量可以唯一地表示成 $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ 的线性组合，这是由于对任何的 s 和 t 有

$$\begin{bmatrix}s\\t\\t\end{bmatrix}=s\begin{bmatrix}1\\0\\1\end{bmatrix}+(t-s)\begin{bmatrix}0\\1\\1\end{bmatrix}+s\begin{bmatrix}0\\1\\0\end{bmatrix}$$

$\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$是 H 的一组基吗？为什么？

26. 在全体实值函数组成的向量空间中，求由$\{\sin t, \sin 2t, \sin t\cos t\}$张成的子空间的一组基．
27. 设 V 是由描述弹簧振子系统振动情况的函数构成的向量空间（参考4.1节习题19）．求 V 的一组基．



习题28的图

28. （RLC 电路）右图中的电路由一个电阻（R 欧姆）、一个感应器（L 亨利）、一个电容器（C 法特）和一个初始电压电源组成．设 $b=R/(2L)$，并且假设选取的 R，L，C 满足 $b=1/\sqrt{LC}$（这是可以做到的，例如，该电路用在电压表中时）．设 $v(t)$（单位：伏特）是在 t 时刻电容器两端的电压，可以证明 v 包含在一个线性变换的零空间 H 中，其中该线性变换将 $v(t)$ 映射到 $Lv''(t)+Rv'(t)+(1/C)v(t)$，H 由所有形如 $v(t)=\mathrm{e}^{-bt}(c_1+c_2t)$的函数组成．试求 H 的基．

习题29和30表明了 $\mathbf{R}^n$的任意一组基一定含有 n 个向量．

29. 设 $S=\{\mathbf{v}_1,\cdots,\mathbf{v}_k\}$是 $\mathbf{R}^n$中 k 个向量的集合，其中 $k<n$．利用1.4节一个定理解释 S 为什么不可能是 $\mathbf{R}^n$的一组基．
30. 设 $S=\{\mathbf{v}_1,\cdots,\mathbf{v}_k\}$是 $\mathbf{R}^n$中 k 个向量的集合，其中 $k>n$．利用第一章的一个定理解释 S 为什么不可能是 $\mathbf{R}^n$的一组基．

习题31和32揭示了线性无关和线性变换之间的一个重要的联系，并且给出了线性相关定义应用的练习. 设 V 和 W 是向量空间，T: $V \to W$ 是一个线性变换，$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$是 V 的子集. 244

31. 证明：如果$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$在 V 中是线性相关的，则像的集合$\{T(\boldsymbol{v}_1),\cdots,T(\boldsymbol{v}_p)\}$在 W 中是线性相关的. 这个事实说明如果一个线性变换将集合$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$映射到一个线性无关集$\{T(\boldsymbol{v}_1),\cdots,T(\boldsymbol{v}_p)\}$，那么原集合也是线性无关的(因为它不可能是线性相关的).
32. 假设 T 是一对一的变换，即如果 $T(\boldsymbol{u})=T(\boldsymbol{v})$，则 $\boldsymbol{u}=\boldsymbol{v}$. 证明：如果像的集合$\{T(\boldsymbol{v}_1),\cdots,T(\boldsymbol{v}_p)\}$是线性相关的，那么$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$也是线性相关的. 这个事实说明在一一的线性变换映射下线性无关集合的像仍然是线性无关的(因为在这种情况下,像的集合不可能是线性相关的).
33. 考虑多项式$\boldsymbol{p}_1(t)=1+t^2$和$\boldsymbol{p}_2(t)=1-t^2$. 在$\mathbf{P}_3$中，$\{\boldsymbol{p}_1,\boldsymbol{p}_2\}$是线性无关的集合吗？为什么？
34. 考虑多项式$\boldsymbol{p}_1(t)=1+t$,$\boldsymbol{p}_2(t)=1-t$和$\boldsymbol{p}_3(t)=2$(对任意的 t). 试写出$\boldsymbol{p}_1$,$\boldsymbol{p}_2$和$\boldsymbol{p}_3$之间的线性相关关系，然后求 Span$\{\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3\}$的一组基.
35. 设 V 是包含线性无关集$\{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3,\boldsymbol{u}_4\}$的向量空间，如何构造 V 中一个向量集合$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4\}$使得$\{\boldsymbol{v}_1,\boldsymbol{v}_3\}$是 Span$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4\}$的基.
36. [**M**]设 $H=\text{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3\}$,$K=\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$，其中

$$\boldsymbol{u}_1=\begin{bmatrix}1\\2\\3\\-1\end{bmatrix},\ \boldsymbol{u}_2=\begin{bmatrix}0\\2\\-1\\1\end{bmatrix},\ \boldsymbol{u}_3=\begin{bmatrix}2\\2\\7\\-3\end{bmatrix}$$

$$\boldsymbol{v}_1=\begin{bmatrix}1\\0\\8\\-4\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}2\\-2\\9\\-5\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}-1\\4\\6\\-2\end{bmatrix}$$

试求 H，K 和 $H+K$ 的基. (见4.1节习题33,34.)
37. [**M**]证明：$\{t,\ \sin t,\ \cos 2t,\ \sin t\cos t\}$是定义在 $\mathbf{R}$ 上的函数的线性无关集. 开始可以设

$$c_1\cdot t+c_2\cdot\sin t+c_3\cdot\cos 2t+c_4\cdot\sin t\cos t=0 \tag{5}$$

对任意的 t 等式(5)都成立，所以可以选择几个特殊值(比如 $t=0,1,2$)得到一个线性方程组，以便确定所有的 c_j 都是零.
38. [**M**]证明：$\{1,\cos t,\cos^2 t,\cdots,\cos^6 t\}$是定义在 $\mathbf{R}$ 上的函数的线性无关集. 利用习题37的方法. (这个结果将在4.5节习题34中用到.)

基础练习答案

1. 设$A=[\boldsymbol{v}_1\ \ \boldsymbol{v}_2]$. 通过行变换可以得到

$$A=\begin{bmatrix}1&-2\\-2&7\\3&-9\end{bmatrix}\sim\begin{bmatrix}1&-2\\0&3\\0&0\end{bmatrix}$$

并不是 A 的每行都含有一个主元位置，所以 A 的列不能张成$\mathbf{R}^3$，由1.4节定理4知. 因此$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$不是$\mathbf{R}^3$的基. 因为$\boldsymbol{v}_1$和$\boldsymbol{v}_2$不在$\mathbf{R}^2$中，它们不可能是$\mathbf{R}^2$的基. 但是，由于$\boldsymbol{v}_1$和$\boldsymbol{v}_2$显然是线性无关的，它们显然是$\mathbf{R}^3$的一个子空间的基，即 Span$\{\boldsymbol{v}_1,\ \boldsymbol{v}_2\}$.
2. 构造矩阵 A，使它的列空间是$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4\}$所张成的空间，然后通过行化简来求 A 的主元列.

$$A=\begin{bmatrix}1&6&2&-4\\-3&2&-2&-8\\4&-1&3&9\end{bmatrix}\sim\begin{bmatrix}1&6&2&-4\\0&20&4&-20\\0&-25&-5&25\end{bmatrix}\sim\begin{bmatrix}1&6&2&-4\\0&5&1&-5\\0&0&0&0\end{bmatrix}$$

245

$\boldsymbol{A}$ 的前两列是主元列，因而构成 Col $\boldsymbol{A}=W$ 的一组基．于是$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$是 W 的一组基．注意，找$\boldsymbol{A}$ 的主元列并不需要知道 $\boldsymbol{A}$ 的简化阶梯矩阵．

3. $\boldsymbol{v}_1$和 $\boldsymbol{v}_2$都不在 H 中，所以$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$不可能是 H 的基．事实上，$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$是所有形如$(c_1,c_2,0)$的向量所构成平面的基，而 H 只是一条直线．

4.4　坐　标　系

给向量空间 V 指定基 $\mathcal{B}$ 的一个重要的原因是希望给 V 设置“坐标系”．本节将要说明，如果 $\mathcal{B}$ 含有 n 个向量，那么坐标系会使 V 看起来像 $\mathbf{R}^n$空间．如果 V 本身就是 $\mathbf{R}^n$，那么 $\mathcal{B}$ 所确定的坐标系将为我们研究 V 提供一个新的“视角”．

坐标系的存在性依赖于下面的基本结论．

【定理7】　唯一表示定理

设 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$是向量空间 V 的一组基，则对于任意 $\boldsymbol{x}\in V$，存在唯一一组数量$c_1,\cdots,c_n$，使得

$$\boldsymbol{x}=c_1\boldsymbol{b}_1+\cdots+c_n\boldsymbol{b}_n \tag{1}$$

证明：由于 $\mathcal{B}$ 张成 V，存在一组数量使得(1)成立．假设还存在数量 $d_1,\cdots,d_n$，使 $\boldsymbol{x}$ 表示成

$$\boldsymbol{x}=d_1\boldsymbol{b}_1+\cdots+d_n\boldsymbol{b}_n$$

那么，两式相减得

$$\mathbf{0}=\boldsymbol{x}-\boldsymbol{x}=(c_1-d_1)\boldsymbol{b}_1+\cdots+(c_n-d_n)\boldsymbol{b}_n \tag{2}$$

因为 $\mathcal{B}$ 是线性无关的，(2)中的权重均为零，即对 $1\leqslant j\leqslant n$ 有 $c_j=d_j$．■

【定义】　假设 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$是 V 的一组基，且 $\boldsymbol{x}\in V$．**$\boldsymbol{x}$ 关于基 $\mathcal{B}$ 的坐标**(coordinates of $\boldsymbol{x}$ relative to the basis $\mathcal{B}$)或**$\boldsymbol{x}$ 的 $\mathcal{B}$－坐标**($\mathcal{B}$－coordinates of $\boldsymbol{x}$)是满足 $\boldsymbol{x}=c_1\boldsymbol{b}_1+\cdots+c_n\boldsymbol{b}_n$ 的权重 $c_1,\cdots,c_n$．

如果 $c_1,\cdots,c_n$ 是 $\boldsymbol{x}$ 的 $\mathcal{B}$－坐标，则 $\mathbf{R}^n$中的向量

$$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}$$

246

就是 $\boldsymbol{x}$(关于 $\mathcal{B}$)的**坐标向量**[coordinate vector of $\boldsymbol{x}$(relative to $\mathcal{B}$)]，或 $\boldsymbol{x}$ 的 **$\mathcal{B}$－坐标向量**($\mathcal{B}$－coordinate vector of $\boldsymbol{x}$)．映射 $\boldsymbol{x}\mapsto[\boldsymbol{x}]_{\mathcal{B}}$是(**由 $\mathcal{B}$ 确定的**)**坐标映射**[coordinate mapping (determined by $\mathcal{B}$)].[1]

【例题1】　考虑 $\mathbf{R}^2$的基 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$，其中 $\boldsymbol{b}_1=\begin{bmatrix}1\\0\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}1\\2\end{bmatrix}$．假设 $\mathbf{R}^2$中的 $\boldsymbol{x}$ 的坐标向量是$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}-2\\3\end{bmatrix}$，求 $\boldsymbol{x}$．

---

1. 坐标映射的概念假设基 $\mathcal{B}$ 是一个加标集，其中的向量以某种预先指定的次序列出．这个性质使得$[\boldsymbol{x}]_{\mathcal{B}}$ 的定义更明确．

解：x 的 $\mathcal{B}$－坐标告诉我们怎样用 $\mathcal{B}$ 中的向量构造 x，即，

$$x=(-2)b_1+3b_2=(-2)\begin{bmatrix}1\\0\end{bmatrix}+3\begin{bmatrix}1\\2\end{bmatrix}=\begin{bmatrix}1\\6\end{bmatrix}$$ ■

【例题 2】 向量 $x=\begin{bmatrix}1\\6\end{bmatrix}$中的分量就是 x 关于标准基 $\varepsilon=\{e_1,e_2\}$的坐标，这是因为

$$\begin{bmatrix}1\\6\end{bmatrix}=1\cdot\begin{bmatrix}1\\0\end{bmatrix}+6\cdot\begin{bmatrix}0\\1\end{bmatrix}=1\cdot e_1+6\cdot e_2$$

于是当 $\varepsilon=\{e_1,e_2\}$时，有$[x]_\varepsilon=x$. ■

4.4.1 坐标的几何解释

一个集合的坐标系就是这个集合中点到 $\mathbf{R}^n$的一对一映射. 例如，在图纸上取定互相垂直的两条轴及每条轴上的单位长度，就构成了平面的一个坐标系. 图 4-15 绘出了标准基$\{e_1, e_2\}$，以及例题 1 中的向量 $b_1(=e_1)$，b_2 和 $x=\begin{bmatrix}1\\6\end{bmatrix}$. 坐标 1 和 6 给出了 x 关于标准基的位置：e_1方向上 1 个单位，e_2方向上 6 个单位.

图 4-16 也绘出了图 4-15 中的向量 b_1，b_2 和 x. 从几何上看，在两幅图中，这 3 个向量都位于一条竖线上. 不过，图 4-16 标准坐标系被抹去，换成了用例题 1 中基 $\mathcal{B}$ 做成的坐标系. 坐标向量$[x]_{\mathcal{B}}=\begin{bmatrix}-2\\3\end{bmatrix}$给出了在新坐标系下 x 的位置：b_1 方向上 -2个单位，b_2 方向上 3 个单位.

247

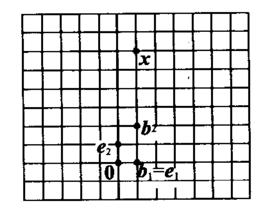


图 4-15 标准坐标纸

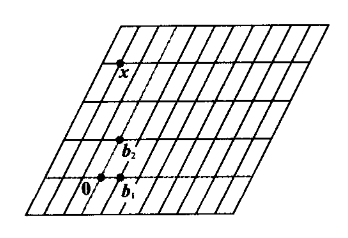


图 4-16 $\mathcal{B}$-坐标纸

【例题 3】 在晶体学中，晶格的描述需要借助 $\mathbf{R}^3$ 中选取的一组基$\{u, v, w\}$，其中 u，v，w 分别对应于晶体中一个单位晶格的三条邻边. 整个晶格由多个单位晶格堆积而成的. 单位晶格有 14 种基本类型，图 4-17 列出了其中三种.[1]

晶体中原子的坐标可以由点阵的基给出. 例如，

$$\begin{bmatrix}1/2\\1/2\\1\end{bmatrix}$$

1. 见 *The Science and Engineering of Materials*, 4th Ed., by Donald R. Askeland (Boston: Prindle, Weber & Schmidt, 2002), p. 36.

就是图4-17b 中晶格上平面中心的原子的坐标.

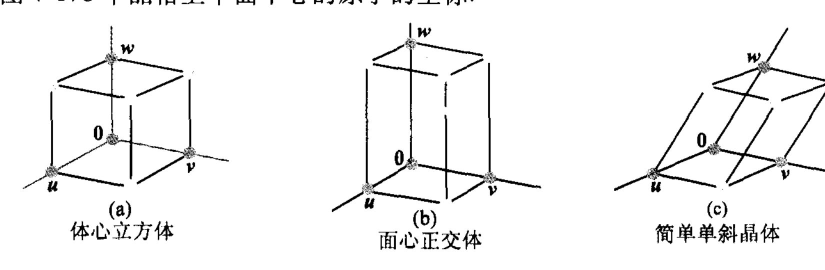


图4-17　单位晶格的例子

4.4.2　$\mathbf{R}^n$中的坐标

固定了$\mathbf{R}^n$的一组基$\mathcal{B}$，我们容易求出已知向量$\boldsymbol{x}$的$\mathcal{B}$-坐标向量. 下面的例题演示了这一算法.

248

【例题4】　设$\boldsymbol{b}_1=\begin{bmatrix}1\\2\end{bmatrix}, \boldsymbol{b}_2=\begin{bmatrix}-1\\1\end{bmatrix}, \boldsymbol{x}=\begin{bmatrix}4\\5\end{bmatrix}$，$\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$. 试求$\boldsymbol{x}$关于$\mathcal{B}$的坐标向量$[\boldsymbol{x}]_{\mathcal{B}}$.

解：$\boldsymbol{x}$的$\mathcal{B}$-坐标c_1，c_2满足

$$c_1\underset{\boldsymbol{b}_1}{\begin{bmatrix}2\\1\end{bmatrix}}+c_2\underset{\boldsymbol{b}_2}{\begin{bmatrix}-1\\1\end{bmatrix}}=\underset{\boldsymbol{x}}{\begin{bmatrix}4\\5\end{bmatrix}}$$

或者

$$\underset{\boldsymbol{b}_1\quad\boldsymbol{b}_2}{\begin{bmatrix}2&-1\\1&1\end{bmatrix}}\begin{bmatrix}c_1\\c_2\end{bmatrix}=\underset{\boldsymbol{x}}{\begin{bmatrix}4\\5\end{bmatrix}}\tag{3}$$

该方程可以通过增广矩阵上的行变换或利用等号左边矩阵的逆来求解. 无论哪种方法，都能得出方程的解为$c_1=3$，$c_2=2$. 因此$\boldsymbol{x}=3\boldsymbol{b}_1+2\boldsymbol{b}_2$，且

$$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}c_1\\c_2\end{bmatrix}=\begin{bmatrix}3\\2\end{bmatrix}$$

见图4-18. ■

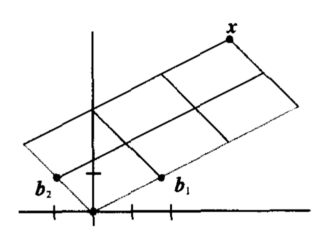


图4-18　$\boldsymbol{x}$的$\mathcal{B}$-坐标向量是(3,2)

(3) 中的矩阵将向量$\boldsymbol{x}$的$\mathcal{B}$-坐标转换成$\boldsymbol{x}$的标准坐标. 在$\mathbf{R}^n$中取定一组基$\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$，也可以进行类似的坐标转换. 设

$$P_{\mathcal{B}}=[\boldsymbol{b}_1 \quad \boldsymbol{b}_2 \quad \cdots \quad \boldsymbol{b}_n]$$

则向量方程

$$\boldsymbol{x}=c_1\boldsymbol{b}_1+c_2\boldsymbol{b}_2+\cdots+c_n\boldsymbol{b}_n$$

等价于

$$\boxed{\boldsymbol{x}=\boldsymbol{P}_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}}} \tag{4}$$

我们称 $\boldsymbol{P}_{\mathcal{B}}$ 是 $\mathbf{R}^n$ 中从 $\mathcal{B}$ 到标准基的**坐标变换矩阵**(change-of-coordinates matrix). 左乘矩阵 $\boldsymbol{P}_{\mathcal{B}}$ 可以将坐标向量 $[\boldsymbol{x}]_{\mathcal{B}}$ 变换为 $\boldsymbol{x}$. 坐标变换方程(4)十分重要，在第 5 章、第 7 章将多处用到.

由于 $\boldsymbol{P}_{\mathcal{B}}$ 的列构成了 $\mathbf{R}^n$ 的一组基，$\boldsymbol{P}_{\mathcal{B}}$ 是可逆的(根据可逆矩阵定理). 左乘 $\boldsymbol{P}_{\mathcal{B}}^{-1}$ 可以将 $\boldsymbol{x}$ 转换到它的 $\mathcal{B}$-坐标向量：

$$\boldsymbol{P}_{\mathcal{B}}^{-1}\boldsymbol{x}=[\boldsymbol{x}]_{\mathcal{B}}$$

这里由 $\boldsymbol{P}_{\mathcal{B}}^{-1}$ 诱导的映射 $\boldsymbol{x}\mapsto[\boldsymbol{x}]_{\mathcal{B}}$ 就是前面提到的坐标映射. 由于 $\boldsymbol{P}_{\mathcal{B}}^{-1}$ 是个可逆矩阵，根据可逆矩阵定理，坐标映射是一个从 $\mathbf{R}^n$ 到 $\mathbf{R}^n$ 的线性变换(见 1.9 节定理 12). 坐标变换的这个性质在拥有基的一般向量空间中也是成立的，我们下面将会看到.

249

4.4.3 坐标映射

给向量空间 V 选定一组基 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 就会诱导出 V 的一个坐标系. 坐标映射 $\boldsymbol{x}\mapsto[\boldsymbol{x}]_{\mathcal{B}}$ 将未知的空间 $\boldsymbol{V}$ 与我们所熟悉的 $\mathbf{R}^n$ 联系起来. 如图 4-19 所示，V 中的点可以由它们的新"名字"所确定.

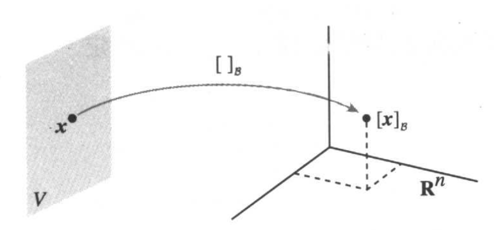


图 4-19 V 到 $\mathbf{R}^n$ 上的坐标映射

【定理 8】 设 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 是向量空间 $\boldsymbol{V}$ 的一组基，则坐标变换 $\boldsymbol{x}\mapsto[\boldsymbol{x}]_{\mathcal{B}}$ 是从 $\boldsymbol{V}$ 到 $\mathbf{R}^n$ 的一对一线性变换.

证明： 任取 V 中的两个向量，比如说，

$$\boldsymbol{u}=c_1\boldsymbol{b}_1+\cdots+c_n\boldsymbol{b}_n$$
$$\boldsymbol{w}=d_1\boldsymbol{b}_1+\cdots+d_n\boldsymbol{b}_n$$

那么，利用向量运算可得

$$\boldsymbol{u}+\boldsymbol{w}=(c_1+d_1)\boldsymbol{b}_1+\cdots+(c_n+d_n)\boldsymbol{b}_n$$

然后得到

$$[\boldsymbol{u}+\boldsymbol{w}]_{\mathcal{B}}=\begin{bmatrix}c_1+d_1\\ \vdots \\ c_n+d_n\end{bmatrix}=\begin{bmatrix}c_1\\ \vdots \\ c_n\end{bmatrix}+\begin{bmatrix}d_1\\ \vdots \\ d_n\end{bmatrix}=[\boldsymbol{u}]_{\mathcal{B}}+[\boldsymbol{w}]_{\mathcal{B}}$$

因此坐标映射保持加法．如果 r 是任一数量，则

$$r\boldsymbol{u} = r(c_1\boldsymbol{b}_1 + \cdots + c_n\boldsymbol{b}_n) = (rc_1)\boldsymbol{b}_1 + \cdots + (rc_n)\boldsymbol{b}_n$$

所以

$$[r\boldsymbol{u}]_{\mathcal{B}} = \begin{bmatrix} rc_1 \\ \vdots \\ rc_n \end{bmatrix} = r\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = r[\boldsymbol{u}]_{\mathcal{B}}$$

250

因此坐标映射也保持数乘，它是一个线性变换．坐标映射是 V 到 $\mathbf{R}^n$ 的一对一映射，其证明见习题 23 和 24． ■

与 1.8 节一样，坐标映射的线性性质可以扩展到线性组合．如果 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p \in V$，$\boldsymbol{c}_1, \cdots, \boldsymbol{c}_p$ 是一组数量，则

$$[c_1\boldsymbol{u}_1 + \cdots + c_p\boldsymbol{u}_p]_{\mathcal{B}} = c_1[\boldsymbol{u}_1]_{\mathcal{B}} + \cdots + c_p[\boldsymbol{u}_p]_{\mathcal{B}} \tag{5}$$

用文字表示，(5)即表明：向量 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p$ 的线性组合的 $\mathcal{B}$－坐标向量就是它们坐标向量的线性组合，两个线性组合的权重完全相同．

定理 8 中的坐标映射是 V 到 $\mathbf{R}^n$ 上同构的一个重要例子．一般地，称一个从向量空间 V 到 W 的一对一的线性变换为 V 到 W 上的**同构**(isomorphism，“iso”源于希腊语中的“相等”，“morph”源于希腊语中的“形式”或“结构”)．V 和 W 的记号和术语可能不同，但它们作为向量空间却是相同的．V 中向量空间上的运算在 W 中依然成立，反之亦然，见习题 25 和 26.

【例题 5】　设 $\mathcal{B}$ 是多项式空间 $\mathbf{P}_3$ 的标准基，即 $\mathcal{B} = \{1, t, t^2, t^3\}$．$\mathbf{P}_3$ 中的任意元素形如

$$\boldsymbol{p}(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$$

由于 $\boldsymbol{p}$ 已经写成了标准基向量的线性组合，我们可以断定

$$[\boldsymbol{p}]_{\mathcal{B}} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

因此坐标映射 $\boldsymbol{p} \mapsto [\boldsymbol{p}]_{\mathcal{B}}$ 是 $\mathbf{P}_3$ 到 $\mathbf{R}^4$ 的同构．$\mathbf{P}_3$ 中所有向量空间上的运算都与 $\mathbf{R}^4$ 中的运算相对应． ■

如果我们想象 $\mathbf{P}_3$ 和 $\mathbf{R}^4$ 分别呈现在两台电脑的屏幕上，电脑之间经由坐标映射来连接，那么，$\mathbf{P}_3$ 所在屏幕上的每个 $\mathbf{P}_3$ 中的向量运算，经过坐标变换，都会以 $\mathbf{R}^4$ 中向量运算的形式准确地重现在另一个屏幕上．$\mathbf{P}_3$ 屏幕上的向量与 $\mathbf{R}^4$ 屏幕上的向量看起来有所不同，但它们的“行为”完全一样，见图 4-20.

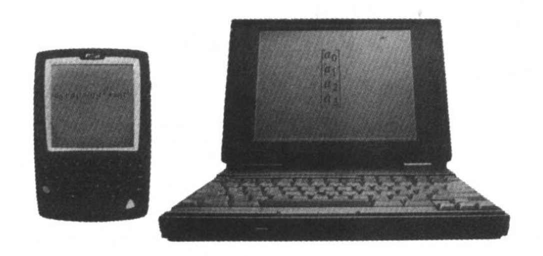


图 4-20　$\mathbf{P}_3$ 空间与 $\mathbf{R}^4$ 同构

【例题 6】 利用坐标向量证明多项式 $1+2t^2, 4+t+5t^2, 3+2t$ 在 $\mathbf{P}_2$ 中线性相关.

解：利用例题 5 的坐标映射可知，上述多项式的坐标向量分别是：(1,0,2)，(4,1,5)和(3,2,0)．将它们写成矩阵 $\boldsymbol{A}$ 的列，通过对 $\boldsymbol{Ax}=\boldsymbol{0}$ 的增广矩阵进行行化简，可以判定它们是否线性无关：

251

$$\begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 2 & 5 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$\boldsymbol{A}$ 的列是线性相关的，因此对应的多项式也是线性相关的．实际上，容易验证 $\boldsymbol{A}$ 的第 3 列等于第 2 列的 2 倍减去第 1 列的 5 倍，对应的多项式关系是

$$3+2t = 2(4+t+5t^2) - 5(1+2t^2)$$ ■

最后一个例题讨论了 $\mathbf{R}^3$ 中同构于 $\mathbf{R}^2$ 的一个平面.

【例题 7】 设 $\boldsymbol{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}, \mathcal{B} = [\boldsymbol{v}_1, \boldsymbol{v}_2]$，$\mathcal{B}$ 是 $H = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$的一组基．判断 $\boldsymbol{x}$ 是否属于 H？如果是，求 $\boldsymbol{x}$ 关于 $\mathcal{B}$ 的坐标向量.

解：如果 $\boldsymbol{x} \in H$，那么下列向量方程是相容的：

$$c_1 \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$$

数量 c_1, c_2 若存在，它们应该是 $\boldsymbol{x}$ 的 $\mathcal{B}-$坐标．利用行变换我们可以得到

$$\begin{bmatrix} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

252

因此 $c_1 = 2, c_2 = 3, [\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$．图 4-21 绘出了由 $\mathcal{B}$ 确定的 H 上的坐标系.

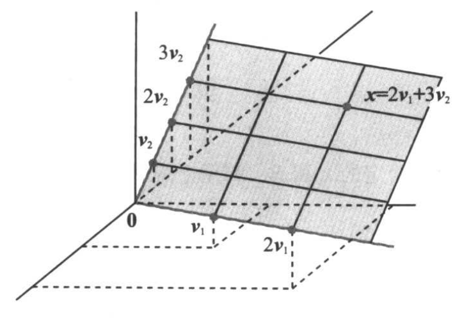


图 4-21 $\mathbf{R}^3$ 中平面 H 上的坐标系

如果选择 H 的另一组基，相应的坐标系是否也会使 H 同构于 $\mathbf{R}^2$？当然，这是一

定的. 我们将在下一节给出证明.

基础练习

1. 设 $\boldsymbol{b}_1=\begin{bmatrix}1\\0\\0\end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix}-3\\4\\0\end{bmatrix}$, $\boldsymbol{b}_3=\begin{bmatrix}3\\-6\\3\end{bmatrix}$, $\boldsymbol{x}=\begin{bmatrix}-8\\2\\3\end{bmatrix}$.

   a. 证明 $\mathcal{B}=\{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3\}$是 $\mathbf{R}^3$ 的一组基.

   b. 求从 $\mathcal{B}$ 到标准基的坐标变换矩阵.

   c. 写出在 $\mathbf{R}^3$ 中 $\boldsymbol{x}$ 与 $[\boldsymbol{x}]_{\mathcal{B}}$ 的关系方程.

   d. 对上面给定的 $\boldsymbol{x}$, 求 $[\boldsymbol{x}]_{\mathcal{B}}$.

2. 集合 $\mathcal{B}=\{1+t, 1+t^2, t+t^2\}$是 $\mathbf{P}_2$ 的一组基. 求 $\boldsymbol{p}(t)=6+3t+t^2$ 关于 $\mathcal{B}$ 的坐标向量.

习题　4.4

在习题 1 ~ 4 中, 求由给定的坐标向量确定的向量 $\boldsymbol{x}$, 并且给出基 $\mathcal{B}$.

1. $\mathcal{B}=\left\{\begin{bmatrix}3\\-5\end{bmatrix},\begin{bmatrix}-4\\6\end{bmatrix}\right\}, [\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}5\\3\end{bmatrix}$

2. $\mathcal{B}=\left\{\begin{bmatrix}4\\5\end{bmatrix},\begin{bmatrix}6\\7\end{bmatrix}\right\}, [\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}8\\-5\end{bmatrix}$

3. $\mathcal{B}=\left\{\begin{bmatrix}1\\-4\\3\end{bmatrix},\begin{bmatrix}5\\2\\-2\end{bmatrix},\begin{bmatrix}4\\-7\\0\end{bmatrix}\right\}, [\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}3\\0\\-1\end{bmatrix}$

4. $\mathcal{B}=\left\{\begin{bmatrix}-1\\2\\0\end{bmatrix},\begin{bmatrix}3\\-5\\2\end{bmatrix},\begin{bmatrix}4\\-7\\3\end{bmatrix}\right\}, [\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}-4\\8\\-7\end{bmatrix}$

253

在习题 5 ~ 8 中, 求 $\boldsymbol{x}$ 关于给定的基 $\mathcal{B}=\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$的坐标向量$[\boldsymbol{x}]_{\mathcal{B}}$.

5. $\boldsymbol{b}_1=\begin{bmatrix}1\\-3\end{bmatrix}, \boldsymbol{b}_2=\begin{bmatrix}2\\-5\end{bmatrix}, \boldsymbol{x}=\begin{bmatrix}-2\\1\end{bmatrix}$

6. $\boldsymbol{b}_1=\begin{bmatrix}1\\-2\end{bmatrix}, \boldsymbol{b}_2=\begin{bmatrix}5\\-6\end{bmatrix}, \boldsymbol{x}=\begin{bmatrix}4\\0\end{bmatrix}$

7. $\boldsymbol{b}_1=\begin{bmatrix}1\\-1\\-3\end{bmatrix}, \boldsymbol{b}_2=\begin{bmatrix}-3\\4\\9\end{bmatrix}, \boldsymbol{b}_3=\begin{bmatrix}2\\-2\\4\end{bmatrix}, \boldsymbol{x}=\begin{bmatrix}8\\-9\\6\end{bmatrix}$

8. $\boldsymbol{b}_1=\begin{bmatrix}1\\0\\3\end{bmatrix}, \boldsymbol{b}_2=\begin{bmatrix}2\\1\\8\end{bmatrix}, \boldsymbol{b}_3=\begin{bmatrix}1\\-1\\2\end{bmatrix}, \boldsymbol{x}=\begin{bmatrix}3\\-5\\4\end{bmatrix}$

在习题 9 和 10 中, 求从 $\mathcal{B}$ 到 $\mathbf{R}^n$ 标准基的坐标变换矩阵.

9. $\mathcal{B}=\left\{\begin{bmatrix}2\\-9\end{bmatrix},\begin{bmatrix}1\\8\end{bmatrix}\right\}$

10. $\mathcal{B}=\left\{\begin{bmatrix}3\\-1\\4\end{bmatrix},\begin{bmatrix}2\\0\\-5\end{bmatrix},\begin{bmatrix}8\\-2\\7\end{bmatrix}\right\}$

在习题 11 和 12 中, 对于给定的 $\boldsymbol{x}$ 和 $\mathcal{B}$, 利用逆矩阵求$[\boldsymbol{x}]_{\mathcal{B}}$.

11. $\mathcal{B}=\left\{\begin{bmatrix}3\\-5\end{bmatrix},\begin{bmatrix}-4\\6\end{bmatrix}\right\}$, $\boldsymbol{x}=\begin{bmatrix}2\\-6\end{bmatrix}$

12. $\mathcal{B}=\left\{\begin{bmatrix}4\\5\end{bmatrix},\begin{bmatrix}6\\7\end{bmatrix}\right\}$, $\boldsymbol{x}=\begin{bmatrix}2\\0\end{bmatrix}$

13. 集合 $\mathcal{B}=\{1+t^2, t+t^2, 1+2t+t^2\}$是 $\mathbf{P}_2$的一组基. 求 $\boldsymbol{p}(t)=1+4t+7t^2$ 关于 $\mathcal{B}$ 的坐标向量.

14. 集合 $\mathcal{B}=\{1-t^2, t-t^2, 2-2t+t^2\}$是 $\mathbf{P}_2$ 的一组基. 求 $\boldsymbol{p}(t)=3+t-6t^2$ 关于 $\mathcal{B}$ 的坐标向量.

在习题 15 和 16 中, 判断各命题的真假, 并说明理由. 除特别说明外, $\mathcal{B}$ 是 V 的一组基.

15. a. 如果 $\boldsymbol{x}$ 在 V 中并且 $\mathcal{B}$ 中有 n 个向量, 则 $\boldsymbol{x}$ 的 $\mathcal{B}$ - 坐标在 $\mathbf{R}^n$ 中.

b. 如果是 P_B 坐标变换矩阵，则对于 V 中的 x 有 $[x]_B = P_B x$.

c. 向量空间 $\mathbf{P}_3$ 与 $\mathbf{R}^3$ 同构.

16. a. 如果 $\mathcal{B}$ 是 $\mathbf{R}^n$ 的标准基，则 x 在 $\mathbf{R}^n$ 中的 $\mathcal{B}$ - 坐标是 x 本身.

b. 称对应 $[x]_B \mapsto x$ 是坐标映射.

c. 在某些情况下，$\mathbf{R}^3$ 中的平面可以同构于 $\mathbf{R}^2$.

17. 向量 $v_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, v_2 = \begin{bmatrix} 2 \\ -8 \end{bmatrix}, v_3 = \begin{bmatrix} -3 \\ 7 \end{bmatrix}$ 张成 $\mathbf{R}^2$，但没有构成它的一组基. 用两种不同的方式将 $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ 表示成 v_1, v_2, v_3 的线性组合.

18. 设 $\mathcal{B} = \{b_1, \cdots, b_n\}$ 是向量空间 V 的一组基，解释 b_1，…，b_n 的 $\mathcal{B}$ - 坐标向量为什么是单位矩阵的列 e_1，…，e_n.

19. 设 S 是向量空间 V 中的一个有限集合，并且 V 中的每个 x 都可以唯一地表示成 S 中元素的线性组合. 证明：S 是 V 的一组基.

20. 假设 $\{v_1, \cdots, v_4\}$ 是向量空间 V 的一个线性相关的张成集. 证明：V 中的每一个 w 表示成 $v_1, \cdots, v_4$ 的线性组合的方式都多于一种. [提示：设 $w = k_1 v_1 + \cdots + k_4 v_4$ 是 V 中的任意向量，利用 $\{v_1, \cdots, v_4\}$ 线性相关性来产生 w 的另一种表示.]

21. 设 $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -4 \end{bmatrix}, \begin{bmatrix} -2 \\ 9 \end{bmatrix} \right\}$ 因为由 $\mathcal{B}$ 确定的坐标变换是一个从 $\mathbf{R}^2$ 到 $\mathbf{R}^2$ 的线性变换，这个变换一定通过某个 2×2 矩阵 A 来实现，试求 A. [提示：通过乘 A 可以将一个向量 x 转化成它的坐标向量 $[x]_B$].

22. 设 $\mathcal{B} = \{b_1, \cdots, b_n\}$ 是 $\mathbf{R}^n$ 的一组基. 求出实现坐标变换 $x \mapsto [x]_B$ 的 $n \times n$ 矩阵 A. [参考习题 21.]

习题 23 ~ 26 是关于向量空间 V，基 $\mathcal{B} = \{b_1, \cdots, b_n\}$ 和坐标变换 $x \mapsto [x]_B$ 的.

23. 证明：坐标变换是一对一的. [提示：假设对 V 中的 u 和 v 有 $[u]_B = [v]_B$，证明 $u = v$.]

24. 证明：坐标映射是到 $\mathbf{R}^n$ 上的，即对 $\mathbf{R}^n$ 中任意给定的 y，其分量为 $y_1, \cdots, y_n$，都存在 V 中的向量 u 使得 $[u]_B = y$.

25. 证明：V 的子集 $\{u_1, \cdots, u_p\}$ 是线性无关的当且仅当坐标向量集 $\{[u_1]_B, \cdots, [u_p]_B\}$ 在 $\mathbf{R}^n$ 中线性无关. [提示：由于坐标映射是一对一的，下列方程有相同的解 $c_1, \cdots, c_p$.]

$c_1 u_1 + \cdots + c_p u_p \quad = \quad \mathbf{0}$　　V 中的零向量

$[c_1 u_1 + \cdots + c_p u_p]_B = [\mathbf{0}]_B$　　$\mathbf{R}^n$ 中的零向量

254

26. 给定 V 中的向量 u 和 w，证明：w 是 $u_1, \cdots, u_p$ 的线性组合当且仅当 $[w]_B$ 是 $[u_1]_B, \cdots, [u_p]_B$ 的线性组合.

在习题 27 ~ 30 中，利用坐标向量判断多项式集合的线性无关性. 说明思路及方法.

27. $1 + t^3, 3 + t - 2t^2, -t + 3t^2 - t^3$

28. $1 - 2t^2 - 3t^3, t + t^3, 1 + 3t - 2t^2$

29. $(t-1)^2, t^3 - 2, (t-2)^3$

30. $(1-t)^3, (2-3t)^2, 3t^2 - 4t^3$

31. 利用坐标向量判断下列多项式集合是否能够张成 $\mathbf{P}_2$，验证你的结论.

a. $1 - 3t + 5t^2, -3 + 5t - 7t^2, -4 + 5t - 6t^2, 1 - t^2$

b. $5t + t^2, 1 - 8t - 2t^2, -3 + 4t + 2t^2, 2 - 3t$

32. 设 $p_1(t)=1+t^2, p_2(t)=2-t+3t^2, p_3(t)=1+2t-4t^2$.

a. 利用坐标向量证明上述向量构成 $\mathbf{P}_2$ 的一组基.

b. 考虑 $\mathbf{P}_2$ 的基 $\mathcal{B}=\{p_1,p_2,p_3\}$，求 $\mathbf{P}_2$ 中的 q，使得 $[q]_{\mathcal{B}}=\begin{bmatrix}-3\\1\\2\end{bmatrix}$.

在习题 33 和 34 中，判断给定的多项式是否是 $\mathbf{P}_3$ 的基，验证你的结论.

33. [M] $3+7t, 5+t-2t^3, t-2t^2, 1+16t-6t^2+2t^3$

34. [M] $5-3t+4t^2+2t^3, 9+t+8t^2-6t^3, 6-2t+5t^2, t^3$

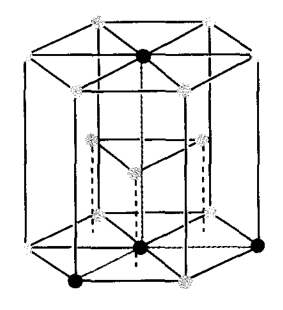
35. [M]设 $H=\text{Span}\{v_1,v_2\}, \mathcal{B}=\{v_1,v_2\}$. 证明 $x\in H$，并且求 x 的 $\mathcal{B}$ - 坐标向量，其中

$$v_1=\begin{bmatrix}11\\-5\\10\\7\end{bmatrix}, v_2=\begin{bmatrix}14\\-8\\13\\10\end{bmatrix}, x=\begin{bmatrix}19\\-13\\18\\15\end{bmatrix}$$

36. [M]设 $H=\text{Span}\{v_1,v_2,v_3\}, \mathcal{B}=\{v_1,v_2,v_3\}$. 证明 $\mathcal{B}$ 是 H 的一组基，$x\in H$ 并且求 x 的 $\mathcal{B}$ - 坐标向量，其中

$$v_1=\begin{bmatrix}-6\\4\\-9\\4\end{bmatrix}, v_2=\begin{bmatrix}8\\-3\\7\\-3\end{bmatrix}, v_3=\begin{bmatrix}-9\\5\\-8\\3\end{bmatrix}, x=\begin{bmatrix}4\\7\\-8\\3\end{bmatrix}$$

[M]习题37 和38 中讨论钛的晶格，这种晶格有六角结构，如4-22 左图所示. $\mathbf{R}^3$ 中的向量 $\begin{bmatrix}2.6\\-1.5\\0\end{bmatrix}, \begin{bmatrix}0\\3\\0\end{bmatrix}, \begin{bmatrix}0\\0\\4.8\end{bmatrix}$ 构成右图单位晶格的一组基，其中数值的单位是埃(Å, 1Å $=10^{-8}$cm).
在钛合金中，钛以外的其他原子可能会包含在*八面体*或*四面体*位置上的单位晶格内(这里所说的八面体和四面体是以这些位置上原子形成的几何形状来命名的).



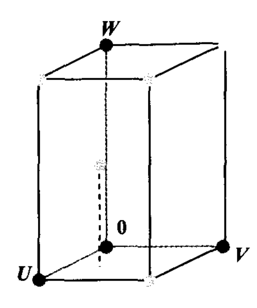


图 4-22 紧凑的六角晶格及其单位晶格

37. 一个八面体关于晶格基的位置是 $\begin{bmatrix}1/2\\1/4\\1/6\end{bmatrix}$，试确定它关于 $\mathbf{R}^3$ 的标准基的坐标.

38. 一个四面体关于晶格基的位置是 $\begin{bmatrix}1/2\\1/2\\1/3\end{bmatrix}$，试确定它关于 $\mathbf{R}^3$ 的标准基的坐标.

基础练习答案

1. a. 显然，矩阵 $\boldsymbol{P}_{\mathcal{B}}=[\boldsymbol{b}_1 \quad \boldsymbol{b}_2 \quad \boldsymbol{b}_3]$行等价于单位矩阵. 由可逆矩阵定理知，$\boldsymbol{P}_{\mathcal{B}}$可逆并且它的列构成 $\mathbf{R}^3$ 的一组基.

255

b. 由(a)得到，坐标转换矩阵是 $\boldsymbol{P}_{\mathcal{B}}=\begin{bmatrix}1&-3&3\\0&4&-6\\0&0&3\end{bmatrix}$.

c. $\boldsymbol{x}=\boldsymbol{P}_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}}$

d. 为解(c)中的方程，通过行化简增广矩阵可能要比直接计算 $\boldsymbol{P}_{\mathcal{B}}^{-1}$更容易些：

$$\underset{\boldsymbol{P}_{\mathcal{B}}\qquad\qquad \boldsymbol{x}}{\begin{bmatrix}1&-3&3&-8\\0&4&-6&2\\0&0&3&3\end{bmatrix}}\sim\underset{\boldsymbol{I}\qquad\qquad [\boldsymbol{x}]_{\mathcal{B}}}{\begin{bmatrix}1&0&0&-5\\0&1&0&2\\0&0&1&1\end{bmatrix}}$$

因此

$$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}-5\\2\\1\end{bmatrix}$$

2. $p(t)=6+3t-t^2$关于 $\mathcal{B}$ 的坐标满足

$$c_1(1+t)+c_2(1+t^2)+c_3(t+t^2)=6+3t-t^2$$

由于 t 的同幂次项系数相同，可以得到

$$\begin{aligned}c_1+c_2&=6\\c_1+c_3&=3\\c_2+c_3&=-1\end{aligned}$$

解得 $c_1=5, c_2=1, c_3=-2$，并且$[\boldsymbol{p}]_{\mathcal{B}}=\begin{bmatrix}5\\1\\-2\end{bmatrix}$.

4.5 向量空间的维数

4.4 节定理 8 表明，如果向量空间 V 的一组基 $\mathcal{B}$ 包含 n 个向量，则 V 同构于 $\mathbf{R}^n$. 本节将要证明：数值 n 是空间 V 的一个本质特征(称作维数)，它不依赖于基的选取. 对维数的讨论将使我们深入理解基的本质.

第一个定理是向量空间 $\mathbf{R}^n$ 中一个著名结论的推广.

【定理 9】 如果向量空间 V 有一组基 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$，则 V 中任何包含多于 n 个向量的集合都是线性相关的.

证明： 设$\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$是 V 中包含多于 n 个向量的集合. 坐标向量$[\boldsymbol{u}_1]_{\mathcal{B}},\cdots,[\boldsymbol{u}_p]_{\mathcal{B}}$构成 $\mathbf{R}^n$中的一个线性相关集合，因为向量的个数(p)大于每个向量中分量的个数(n). 因此存在不全为零的数量 $c_1,\cdots,c_p$，使得

256

$$c_1[\boldsymbol{u}_1]_{\mathcal{B}}+\cdots+c_p[\boldsymbol{u}_p]_{\mathcal{B}}=\begin{bmatrix}0\\\vdots\\0\end{bmatrix}, \mathbf{R}^n\text{中的零向量}$$

由于坐标映射是线性变换，我们有

$$[c_1\boldsymbol{u}_1+\cdots+c_p\boldsymbol{u}_p]_{\mathcal{B}}=\begin{bmatrix}0\\ \vdots\\ 0\end{bmatrix}$$

等式右端的零向量中包含了由 $\mathcal{B}$ 中基向量构造 $c_1\boldsymbol{u}_1+\cdots+c_p\boldsymbol{u}_p$ 所需的 n 个权重. 即 $c_1\boldsymbol{u}_1+\cdots+c_p\boldsymbol{u}_p=0\cdot\boldsymbol{b}_1+\cdots+0\cdot\boldsymbol{b}_n=\boldsymbol{0}$. 由于 c_i 不全为零，因此 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是线性相关的. ■

定理9表明，如果向量空间 V 有一组基 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$，则 V 中的每个线性无关集都不能包含多于 n 个向量.[1]

【定理10】 如果向量空间 V 有一组由 n 个向量构成的基，则 V 的每组基恰好包含 n 个向量.

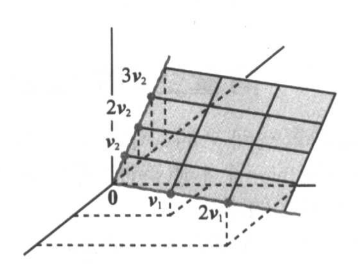
证明： 设 $\mathcal{B}_1$ 是一组含 n 个向量的基，$\mathcal{B}_2$ 是其他任意一组基. 由于 $\mathcal{B}_1$ 是基且 $\mathcal{B}_2$ 线性无关，则由定理9可知，$\mathcal{B}_2$ 中不能多于 n 个向量. 又因为 $\mathcal{B}_2$ 是基且 $\mathcal{B}_1$ 线性无关，$\mathcal{B}_2$ 中至少有 n 个向量. 因此 $\mathcal{B}_2$ 恰含有 n 个向量. ■

如果非零向量空间 V 是由有限集合 S 张成的，则由张成集定理知，存在 S 的一个子集作成 V 的基. 在这种情形下，定理10保证了下述定义合理.

【定义】 如果 V 由有限集张成，则称 V 是**有限维的**(finite-dimensional)，V 的**维数**(dimensional)，记作 dim V，是 V 的基所含向量的个数. 零向量空间 $\{\boldsymbol{0}\}$ 的维数定义为零. 如果 V 不是由有限集张成的，就称 V 是**无限维的**(infinite-dimensional).

【例题1】 $\mathbf{R}^n$ 的标准基含有 n 个向量，所以 $\dim\mathbf{R}^n=n$. 标准多项式基 $\{1,t,t^2\}$ 表明 $\dim\mathbf{P}_2=3$. 一般地，有 $\dim\mathbf{P}_n=n+1$. 全体多项式构成的空间 $\mathbf{P}$ 是无限维的(习题27). ■

257



【例题2】 设 $H=\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$，其中 $\boldsymbol{v}_1=\begin{bmatrix}3\\6\\2\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}-1\\0\\1\end{bmatrix}$，则 H 是4.4节例题7所讨论的平面. $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是 H 的一组基，这是因为 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 之间没有倍数关系，两者线性

1. 定理9也可以应用到 V 的无限集中. 称一个无限集是线性相关的如果它的某个有限子集是线性相关的，否则，就是线性无关的. 如果 S 是 V 中的一个无限子集，取 S 的任意子集 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$，其中 $p>n$，上面的证明说明该子集是线性相关的，因此 S 也是线性相关的.

无关. 因此 $\dim H=2$. ■

【例题 3】 求子空间 H 的维数，其中

$$H=\left\{\begin{bmatrix}a-3b+6c\\5a+4d\\b-2c-d\\5d\end{bmatrix}: a,b,c,d\in\mathbf{R}\right\}$$

解：易见 H 是下列向量的全体线性组合所构成的集合：

$$\boldsymbol{v}_1=\begin{bmatrix}1\\5\\0\\0\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}-3\\0\\1\\0\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}6\\0\\-2\\0\end{bmatrix},\boldsymbol{v}_4=\begin{bmatrix}0\\4\\-1\\5\end{bmatrix}$$

显然，$\boldsymbol{v}_1\neq\mathbf{0}$，$\boldsymbol{v}_2$不是$\boldsymbol{v}_1$的倍数，但$\boldsymbol{v}_3$是$\boldsymbol{v}_2$的倍数. 由张成集定理可知，去除$\boldsymbol{v}_3$后的集合仍然可以张成 H. 最后，$\boldsymbol{v}_4$不是$\boldsymbol{v}_1$和$\boldsymbol{v}_2$的线性组合，所以$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_4\}$是线性无关的（由 4.3 节定理 4），进而是 H 的一组基. 所以 $\dim H=3$. ■

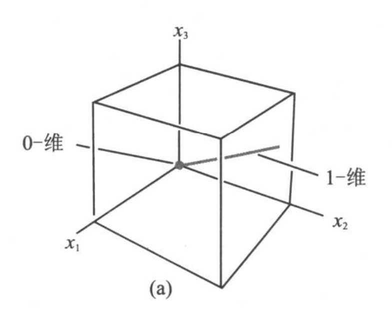
【例题 4】 $\mathbf{R}^3$的子空间可以根据维数来分类，见图 4-23.

0 维子空间：只有零空间.

1 维子空间：由单个非零向量张成的任意子空间，这样的子空间是过原点的直线.

2 维子空间：两个线性无关的非零向量张成的任意子空间，这样的子空间是过原点的平面.

3 维子空间：只有 $\mathbf{R}^3$本身，根据可逆矩阵定理，$\mathbf{R}^3$中任意 3 个线性无关的向量都能张成 $\mathbf{R}^3$.



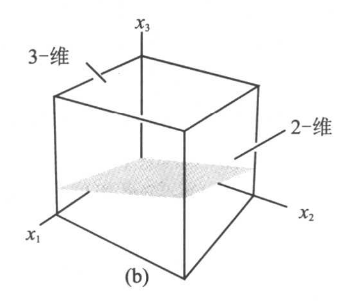


图 4-23 $\mathbf{R}^3$子空间的例子

258

4.5.1 有限维空间的子空间

下面的定理与张成集定理互为补充.

【定理 11】 设 H 是有限维向量空间 V 的子空间. H 中的任何线性无关集都可以扩张成 H 的基. 此外，H 是有限维的，并且 $\dim H\leqslant\dim V$.

证明： 如果 $H=\{\mathbf{0}\}$，则一定有 $\dim H=0\leqslant\dim V$. 否则，设 $S=\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_k\}$是 H 的任意线性无关集. 如果 S 张成 H，则 S 是 H 的一组基. 否则，存在 H 的某个 $\boldsymbol{u}_{k+1}$不属于 Span S. $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_k,\boldsymbol{u}_{k+1}\}$将是线性无关的，因为该集合中的向量都不能表示成其他向量的线性组合(根据定理4).

如果新集合不能张成 H，我们就继续这个过程，将 S 扩张成 H 的极大线性无关集. 但由定理9可知，在 S 的线性无关扩张中，向量个数不能超过 V 的维数. 所以 S 的扩张最终能张成 H，成为 H 的一组基，并且 $\dim H\leqslant\dim V$. ■

如果已知向量空间或者子空间的维数，那么下面的定理可以简化求基的过程. 即，如果集合中元素数目刚好合适，那么只需要证明该集合是线性无关的，或者它可以张成整个空间. 这个定理在许多实际应用(例如,涉及微分方程或差分方程的问题)中非常重要，在这些应用中验证线性无关性要比验证张成性容易得多.

【定理12】 基定理

设 V 是一个 p 维向量空间，$p\geqslant 1$. V 中任何 p 元线性无关集都是 V 的一组基. 任何可以张成 V 的 p 元集合都是 V 的一组基.

证明： 根据定理11，含 p 个元素的线性无关集可以扩张成 V 的基. 但由于 $\dim V=p$，这个基必须只含有 p 个元素，所以 S 本身就是 V 的基. 现在假设 S 含有 p 个元素并且张成 V，由于 V 是非零的，由张成集定理可知，S 的一个子集 S' 是 V 的基. 因为 $\dim V=p$，S'必须包含 p 个向量，因此 $S=S'$. ■

4.5.2 Nul A 和 Col A 的维数

由于矩阵 $\boldsymbol{A}$ 的主元列构成 Col $\boldsymbol{A}$ 的一组基，因此一旦知道主元列，就知道了 Col $\boldsymbol{A}$的维数. 求 Nul $\boldsymbol{A}$ 的维数看起来很困难，因为求 Nul $\boldsymbol{A}$ 的基通常比求Col $\boldsymbol{A}$的基要复杂. 不过，这里却有个捷径！

259

令 $\boldsymbol{A}$ 是一个 $m\times n$ 矩阵，在4.2节中我们学习了求 Nul $\boldsymbol{A}$ 的张成集的标准方法. 假设方程 $\boldsymbol{Ax}=\mathbf{0}$ 有 k 个自由变量. 它们恰好生成 k 个线性无关的向量，比如说 $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_k$，其中每一个都对应一个自由变量. 所以$\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_k\}$是 Nul $\boldsymbol{A}$ 的基，且自由变量的个数决定了基的大小. 我们将这些事实总结如下，供下文参考.

Nul $\boldsymbol{A}$ 的维数是方程 $\boldsymbol{Ax}=\mathbf{0}$ 中自由变量的个数，Col $\boldsymbol{A}$ 的维数是 $\boldsymbol{A}$ 中主元列的个数.

【例题5】 求 $\boldsymbol{A}$ 的零空间和列空间的维数，其中

$$\boldsymbol{A}=\begin{bmatrix}-3 & 6 & -1 & 1 & -7\\ 1 & -2 & 2 & 3 & -1\\ 2 & -4 & 5 & 8 & -4\end{bmatrix}$$

解： 通过行化简将增广矩阵$[\boldsymbol{A}\quad\mathbf{0}]$化成阶梯形式：

$$\begin{bmatrix}1 & -2 & 2 & 3 & -1 & 0\\ 0 & 0 & 1 & 2 & -2 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

自由变量有3个：x_2,x_4和 x_5，因此 Nul $\boldsymbol{A}$ 的维数是3. 此外，因为 $\boldsymbol{A}$ 有两个主元列，所以 $\dim\text{Col}\,\boldsymbol{A}=2$. ■

基础练习

判断各命题的真假，并说明理由. 这里 V 是一个非零的有限维向量空间.

1. 如果 $\dim V=p$，S 是 V 的线性相关子集，则 S 中包含的向量多于 p 个.
2. 如果 S 张成 V，T 是 V 的子集，并且 T 中的向量多于 S，则 T 线性相关.

习题 4.5

对习题 1 ~ 8 中的每个子集，(a)求它们一组基，(b)求出它们的维数.

1. $\left\{\begin{bmatrix} s-2t \\ s+t \\ 3t \end{bmatrix}: s,t\in\mathbf{R}\right\}$ 2. $\left\{\begin{bmatrix} 4s \\ -3s \\ -t \end{bmatrix}: s,t\in\mathbf{R}\right\}$ 3. $\left\{\begin{bmatrix} 2c \\ a-b \\ b-3c \\ a+2b \end{bmatrix}: a,b,c\in\mathbf{R}\right\}$

4. $\left\{\begin{bmatrix} a+b \\ 2a \\ 3a-b \\ -b \end{bmatrix}: a,b\in\mathbf{R}\right\}$ 5. $\left\{\begin{bmatrix} a-4b-2c \\ 2a+5b-4c \\ -a+2c \\ -3a+7b+6c \end{bmatrix}: a,b,c\in\mathbf{R}\right\}$ 6. $\left\{\begin{bmatrix} 3a+6b-c \\ 6a-2b-2c \\ -9a+5b+3c \\ -3a+b+c \end{bmatrix}: a,b,c\in\mathbf{R}\right\}$ 260

7. $\{(a,b,c): a-3b+c=0, b-2c=0, 2b-c=0\}$
8. $\{(a,b,c,d): a-3b+c=0\}$
9. 求 $\mathbf{R}^3$ 中全体第 1 个和第 3 个分量相同的向量构成的子空间的维数.
10. 求 $\mathbf{R}^2$ 中由 $\begin{bmatrix} 2 \\ -5 \end{bmatrix}$，$\begin{bmatrix} -4 \\ 10 \end{bmatrix}$，$\begin{bmatrix} -3 \\ 6 \end{bmatrix}$ 张成的子空间的维数.

在习题 11 和 12 中，求由给定的向量张成的子空间的维数.

11. $\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix},\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix},\begin{bmatrix} 9 \\ 4 \\ -2 \end{bmatrix},\begin{bmatrix} -7 \\ -3 \\ 1 \end{bmatrix}$ 12. $\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix},\begin{bmatrix} -3 \\ 4 \\ 1 \end{bmatrix},\begin{bmatrix} -8 \\ 6 \\ 5 \end{bmatrix},\begin{bmatrix} -3 \\ 0 \\ 7 \end{bmatrix}$

对习题 13 ~ 18 中给出的矩阵，确定 Nul $\boldsymbol{A}$ 和 Col $\boldsymbol{A}$ 的维数.

13. $\boldsymbol{A}=\begin{bmatrix} 1 & -6 & 9 & 0 & -2 \\ 0 & 1 & 2 & -4 & 5 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ 14. $\boldsymbol{A}=\begin{bmatrix} 1 & 3 & -4 & 2 & -1 & 6 \\ 0 & 0 & 1 & -3 & 7 & 0 \\ 0 & 0 & 0 & 1 & 4 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

15. $\boldsymbol{A}=\begin{bmatrix} 1 & 0 & 9 & 5 \\ 0 & 0 & 1 & -4 \end{bmatrix}$ 16. $\boldsymbol{A}=\begin{bmatrix} 3 & 4 \\ -6 & 10 \end{bmatrix}$ 17. $\boldsymbol{A}=\begin{bmatrix} 1 & -1 & 0 \\ 0 & 4 & 7 \\ 0 & 0 & 5 \end{bmatrix}$ 18. $\boldsymbol{A}=\begin{bmatrix} 1 & 4 & -1 \\ 0 & 7 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

在习题 19 和 20 中，V 是一个向量空间. 判断各命题的真假，并说明理由.

19. a. 矩阵的主元列个数等于列空间的维数.
    b. $\mathbf{R}^3$ 中的平面是 $\mathbf{R}^3$ 的 2 维子空间.
    c. 向量空间 $\mathbf{P}_4$ 的维数是 4.
    d. 如果 $\dim V=n$，并且 S 是 V 的线性无关集，则 S 是 V 的基.
    e. 如果 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 张成一个有限维向量空间 V，T 是 V 中一个多于 p 个向量的集合，则 T

是线性相关的.

20. a. $\mathbf{R}^2$是$\mathbf{R}^3$的2维子空间.
 b. 方程$A\boldsymbol{x}=\mathbf{0}$中变量的个数等于Nul A的维数.
 c. 一个向量空间是无限维的，如果它是由一个无限集合张成的.
 d. 如果$\dim V=n$，并且S张成V，则S是V的基.
 e. $\mathbf{R}^3$的3维子空间只有它本身.
21. 前四个埃尔米特多项式是1，$2t$，$-2+4t^2$和$-12t+8t^3$，这些多项式来自数学物理学中的某些重要的微分方程的研究[1]，证明：前四个埃尔米特多项式构成$\mathbf{P}_3$的一组基.
22. 前四个拉盖尔多项式是1，$1-t$，$2-4t+t^2$和$6-18t+9t^2-t^3$，证明：这些多项式构成$\mathbf{P}_3$的一组基.
23. 设$\mathcal{B}$是$\mathbf{P}_3$的一组基，并且包含习题21中的埃尔米特多项式，$\boldsymbol{p}(t)=7-12t-8t^2+12t^3$. 求$\boldsymbol{p}$关于$\mathcal{B}$的向量坐标.
24. 设$\mathcal{B}$是$\mathbf{P}_2$的一组基，并且包含习题22中的列出的拉盖尔多项式，$\boldsymbol{p}(t)=7-8t+3t^2$,求$\boldsymbol{p}$关于$\mathcal{B}$的向量坐标.
25. 设S是n维向量空间V的一个子集，假设S中向量的个数小于n，解释S为什么不能张成V.
26. 设H是n维向量空间V的一个n维子空间，证明$H=V$.

261

27. 解释为什么所有多项式构成的空间是无限维空间.
28. 证明：定义在实数轴上的所有连续函数构成的空间$C(\mathbf{R})$是无限维空间.

在习题29和30中，V是一个非零的有限维向量空间，所有列出的向量属于V，判断各命题的真假，并说明理由.（这些问题较习题19和20更难.）

29. a. 如果存在一个集合$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$张成$V$，则$\dim V\leqslant p$.
 b. 如果V中存在一个线性相关集合$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$，则$\dim V\geqslant p$.
 c. 如果$\dim V=p$，则V中存在一个含$p+1$个向量的张成集.
30. a. 如果V中存在一个线性相关集合$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$，则$\dim V\leqslant p$.
 b. 如果V中的每个p元集合都不能张成V，则$\dim V>p$.
 c. 如果$p\geqslant 2$并且$\dim V=p$，则每个$p-1$元子集是线性无关的.

习题31和32中讨论的是有限维向量空间V和W以及线性变换T：$V\mapsto W$.

31. 设H是V的非零子空间，$T(H)$是H中向量像的集合，则由4.2节习题35可知，$T(H)$是W的一个子空间. 证明：$\dim T(H)\leqslant\dim H$.
32. 设H是V的非零子空间，T是V到W内的一对一(线性)映射. 证明：$\dim T(H)=\dim H$. 如果T恰好是V到W上的一对一(线性)映射，则$\dim V=\dim W$. 同构的无限维向量空间具有相同的维数.
33. [M]根据定理11，$\mathbf{R}^n$中的线性无关集$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k\}$可以扩张成$\mathbf{R}^n$的一组基. 一个方法就是构造$A=[\boldsymbol{v}_1\cdots\boldsymbol{v}_k\quad\boldsymbol{e}_1\cdots\boldsymbol{e}_n]$，其中$\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n$是单位矩阵的列，则$A$的主元列构成$\mathbf{R}^n$的一组基.
 a. 利用上面介绍的方法来将下列向量扩充成$\mathbf{R}^5$的基：

$$\boldsymbol{v}_1=\begin{bmatrix}-9\\-7\\8\\-5\\7\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}9\\4\\1\\6\\-7\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}6\\7\\-8\\5\\-7\end{bmatrix}$$

1. 见*Introduction to Functional Analysis*, 2d ed., by A. E. Taylor and David C. Lay (New York: John Wiley & Sons, 1980), pp. 92－93. 该书还讨论了其他多项式族.

b. 解释为什么该方法在一般情况下也适用：为什么初始向量 $v_1, \cdots, v_k$ 包含在所求的 Col A 的基中？为什么 Col $A = \mathbf{R}^n$？

34. [M]设 $\mathcal{B} = \{1, \cos t, \cos^2 t, \cdots, \cos^6 t\}$，$\mathcal{C} = \{1, \cos t, \cos 2t, \cdots, \cos 6t\}$. 假定下列三角恒等式成立(见4.1节习题37).

$\cos 2t = -1 + 2\cos^2 t$

$\cos 3t = -3\cos t + 4\cos^3 t$

$\cos 4t = 1 - 8\cos^2 t + 8\cos^4 t$

$\cos 5t = 5\cos t - 20\cos^3 t + 16\cos^5 t$

$\cos 6t = -1 + 18\cos^2 t - 48\cos^4 t + 32\cos^6 t$

设 H 是由 $\mathcal{B}$ 中的函数张成的函数子空间，则由4.3节习题38知，$\mathcal{B}$ 是 H 的基.

a. 写出 $\mathcal{C}$ 中向量的 $\mathcal{B}$-坐标向量，并利用它们证明：$\mathcal{C}$ 是 H 的线性无关集.

b. 解释 $\mathcal{C}$ 为什么是 H 的基.

基础练习答案

1. 错，考虑集合 $\{\mathbf{0}\}$.
2. 对，根据张成集定理，S 包含 V 的一组基，称它为 S'. 则 T 中的向量多于 S'. 由定理9可知，T 是线性相关的.

4.6 秩

本节将借助向量空间的概念，观察矩阵内部构造，进而揭示矩阵行、列之间一些十分有趣且实用的联系.

例如，设想一下，将2000个随机数放入 40×50 矩阵 A 中，判断 A 中线性无关列的最大个数和 A^T 中线性无关列(A 中的行)的最大个数. 注意，这两个数是一样的. 我们将要看到，这个公共的值就是矩阵的秩. 为了解释其中的缘由，我们需要考察由 A 的行张成的子空间. 262

4.6.1 行空间

如果 A 是个 $m \times n$ 矩阵，A 的每一行有 n 个分量，因此可以将其等同于 $\mathbf{R}^n$ 中的向量. 称行向量的全体线性组合构成的集合为 A 的**行空间**(row space)，记作 Row A. 由于每行都有 n 个分量，Row A 是 $\mathbf{R}^n$ 的一个子空间. 因为 A 的行就是 A^T 中的列，我们可以用 Col A^T 来代替 Row A.

【例题1】 设

$$A = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix} \quad \text{且} \quad \begin{aligned} \boldsymbol{r}_1 &= (-2, -5, 8, 0, -17) \\ \boldsymbol{r}_2 &= (1, 3, -5, 1, 5) \\ \boldsymbol{r}_3 &= (3, 11, -19, 7, 1) \\ \boldsymbol{r}_4 &= (1, 7, -13, 5, -3) \end{aligned}$$

A 的行空间是 $\mathbf{R}^5$ 中由 $\{\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \boldsymbol{r}_4\}$ 张成的子空间，即 Row A = Span$\{\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \boldsymbol{r}_4\}$. 将行向量横着写是十分自然的，但是为方便起见，我们也可以将它竖着写，即写成列向量. ■

在例题1中，如果我们知道 A 各行之间的线性相关关系，就可以利用张成集定理来把张成集缩小成基. 遗憾的是，A 上的行变换不能给出这一信息，因为行变换会改变各行之间的相关关系. 但是，正如下面的定理所言，对 A 进行行化简是非常有用的！

【定理 13】　如果两个矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是行等价的，则它们的行空间相等．如果 $\boldsymbol{B}$ 是阶梯形矩阵，则 $\boldsymbol{B}$ 的非零行构成 $\boldsymbol{B}$ 的行空间的一组基，也构成 $\boldsymbol{A}$ 的行空间的一组基.

证明： 如果 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 通过行变换得到的，那么 $\boldsymbol{B}$ 中的每行都是 $\boldsymbol{A}$ 中行的线性组合．这就说明，$\boldsymbol{B}$ 中行的线性组合自然也是 $\boldsymbol{A}$ 中行的线性组合．因此 $\boldsymbol{B}$ 的行空间就包含在 $\boldsymbol{A}$ 的行空间中．由于行变换是可逆的，类似地，可以证明 $\boldsymbol{A}$ 的行空间是 $\boldsymbol{B}$ 的行空间的一个子集．所以，这两个行空间相等．如果 $\boldsymbol{B}$ 是阶梯形矩阵，那么 $\boldsymbol{B}$ 中任意非零行都不是其后非零行的线性组合，因此 $\boldsymbol{B}$ 的全体非零行线性无关(将 $\boldsymbol{B}$ 的非零行逆序排列，第1行排在最后，然后利用定理4)．因此 $\boldsymbol{B}$ 的非零行构成 $\boldsymbol{B}$ 的和 $\boldsymbol{A}$ 的(公共)行空间的一组基. ■

263

本节的主要结论涉及以下三个空间：Row $\boldsymbol{A}$，Col $\boldsymbol{A}$ 和 Nul $\boldsymbol{A}$．下面的例题为证明后面的主要结论做准备，它演示了怎样利用 $\boldsymbol{A}$ 上的一个行变换序列求这三个空间的基.

【例题 2】　分别求下列矩阵行空间、列空间和零空间的基.

$$\boldsymbol{A}=\begin{bmatrix}-2 & -5 & 8 & 0 & -17\\ 1 & 3 & -5 & 1 & 5\\ 3 & 11 & -19 & 7 & 1\\ 1 & 7 & -13 & 5 & -3\end{bmatrix}$$

解： 为求行空间和列空间的基，通过行化简将 $\boldsymbol{A}$ 化成阶梯形式：

$$\boldsymbol{A}\sim\boldsymbol{B}=\begin{bmatrix}1 & 3 & -5 & 1 & 5\\ 0 & 1 & -2 & 2 & -7\\ 0 & 0 & 0 & -4 & 20\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

根据定理13，$\boldsymbol{B}$ 的前3行构成 $\boldsymbol{A}$ 的行空间(同样也是 $\boldsymbol{B}$ 的行空间)的一组基，因此

$$\text{Row}\,\boldsymbol{A}\text{ 的基：}\{(1,3,-5,1,5),(0,1,-2,2,-7),(0,0,0,-4,20)\}$$

对于列空间，观察 $\boldsymbol{B}$ 可知其主元列是第1、2和4列．因此 $\boldsymbol{A}$(非 $\boldsymbol{B}$)的第1、2和4列构成 Col $\boldsymbol{A}$ 的一组基：

$$\text{Col}\,\boldsymbol{A}\text{ 的基：}\left\{\begin{bmatrix}-2\\ 1\\ 3\\ 1\end{bmatrix},\begin{bmatrix}-5\\ 3\\ 11\\ 7\end{bmatrix},\begin{bmatrix}0\\ 1\\ 7\\ 5\end{bmatrix}\right\}$$

注意，$\boldsymbol{A}$ 的任意阶梯形式都给出了 Row $\boldsymbol{A}$ 的一组基(这组基实际上由阶梯形矩阵的非零行构成)对于 Col $\boldsymbol{A}$，$\boldsymbol{A}$ 的阶梯形式指明了 $\boldsymbol{A}$ 的主元列，进而给出了 Col $\boldsymbol{A}$ 的一组基．但是对于 Nul $\boldsymbol{A}$，我们需要利用简化阶梯形式．对 $\boldsymbol{B}$ 继续施行行变换得到

$$\boldsymbol{A}\sim\boldsymbol{B}\sim\boldsymbol{C}=\begin{bmatrix}1 & 0 & 1 & 0 & 1\\ 0 & 1 & -2 & 0 & 3\\ 0 & 0 & 0 & 1 & -5\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

方程 $\boldsymbol{Ax}=\boldsymbol{0}$ 等价于 $\boldsymbol{Cx}=\boldsymbol{0}$，即，

$$\begin{aligned} x_1 + x_3 + x_5 &= 0 \\ x_2 - 2x_3 + 3x_5 &= 0 \\ x_4 - 5x_5 &= 0 \end{aligned}$$

264

所以 $x_1=-x_3-x_5, x_2=2x_3-3x_5, x_4=5x_5$，其中 x_3 和 x_5 是自由变量．由计算可得

$$\text{Nul } \boldsymbol{A} \text{ 的基：} \left\{ \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \\ 0 \\ 5 \\ 1 \end{bmatrix} \right\}$$

观察所求的三组基可以发现，与 Col $\boldsymbol{A}$ 的基不同，Row $\boldsymbol{A}$ 和 Nul $\boldsymbol{A}$ 的基与 $\boldsymbol{A}$ 的元素没有显而易见的联系[1]． ■

警告：尽管例题 2 中 $\boldsymbol{B}$ 的前 3 行是线性无关的，但由此断定 $\boldsymbol{A}$ 的前 3 行也线性无关是错误的(实际上，$\boldsymbol{A}$ 的第 3 行等于第 1 行乘以 2 加上第 2 行乘以 7)．行变换不保持矩阵行之间的线性相关关系．

4.6.2 秩定理

下面的定理描述了 Col $\boldsymbol{A}$，Row $\boldsymbol{A}$ 和 Nul $\boldsymbol{A}$ 三者维数之间的基本关系．

【定义】 $\boldsymbol{A}$ 的**秩**(rank)是 $\boldsymbol{A}$ 的列空间的维数．

由于 Row $\boldsymbol{A}$ 与 Col $\boldsymbol{A}^{\mathrm{T}}$ 相同，$\boldsymbol{A}$ 的行空间的维数就是 $\boldsymbol{A}^{\mathrm{T}}$ 的秩．$\boldsymbol{A}$ 的零空间的维数有时被称为 $\boldsymbol{A}$ 的**零化度**(nullity)，不过本书不使用这一术语．

细心的读者可能在完成 4.5 节习题或者阅读上面的例题 2 时，就已经发现了下面定理的部分或全部结论．

【定理 14】 秩定理

$m\times n$ 矩阵 $\boldsymbol{A}$ 的行空间和列空间有相同的维数．这个公共的维数，即 $\boldsymbol{A}$ 的秩，也等于 $\boldsymbol{A}$ 中主元位置的个数并且它满足等式

$$\text{rank } \boldsymbol{A} + \dim \text{Nul } \boldsymbol{A} = n$$

证明：根据 4.3 节定理 6，rank $\boldsymbol{A}$ 等于 $\boldsymbol{A}$ 中主元列的个数．或者说，rank $\boldsymbol{A}$ 等于 $\boldsymbol{A}$ 的阶梯形式 $\boldsymbol{B}$ 当中的主元位置的个数．此外，由于对每个主元，$\boldsymbol{B}$ 中都有一个非零行与之对应，这些行构成 $\boldsymbol{A}$ 的行空间的一组基，因此 $\boldsymbol{A}$ 的秩也等于其行空间的维数．

265

由 4.5 节可知，Nul $\boldsymbol{A}$ 的维数等于方程 $\boldsymbol{Ax}=\boldsymbol{0}$ 中自由变量的个数．换言之，Nul $\boldsymbol{A}$ 的维数就是 $\boldsymbol{A}$ 中非主元列的个数(与 Nul $\boldsymbol{A}$ 有关联的是这些列的数目，而不是列本身)．显然，

$$\{\text{主元列的个数}\} + \{\text{非主元列的个数}\} = \{\text{列的个数}\}$$

定理得证． ■

定理 14 所蕴含的思想在例题 2 的演算过程中已有体现：阶梯形式 $\boldsymbol{B}$ 中的 3 个主

---

1. 可以找到 Row $\boldsymbol{A}$ 的一组基，其中的向量都是 $\boldsymbol{A}$ 的行：首先写出 $\boldsymbol{A}^{\mathrm{T}}$，然后对 $\boldsymbol{A}^{\mathrm{T}}$ 进行行化简直到找出它的主元列．$\boldsymbol{A}^{\mathrm{T}}$ 的这些主元列都是 $\boldsymbol{A}$ 的行，因此它们构成 $\boldsymbol{A}$ 的行空间的一组基．

元位置确定了基本变量，也确定了 Col $\boldsymbol{A}$ 和 Row $\boldsymbol{A}$ 的基向量.

【例题 3】

a. 如果 $\boldsymbol{A}$ 是 7×9 矩阵并且有 2 维零空间，则 $\boldsymbol{A}$ 的秩是多少？

b. 一个 6×9 矩阵可以有 2 维零空间吗？

解：

a. 由于 $\boldsymbol{A}$ 有 9 列，$(\text{rank}\,\boldsymbol{A})+2=9$，因此 $\text{rank}\,\boldsymbol{A}=7$.

b. 不可以. 如果一个 6×9 矩阵 $\boldsymbol{B}$ 有 2 维零空间，根据秩定理，$\boldsymbol{B}$ 的秩必须是 7. 但是，$\boldsymbol{B}$ 的列是 $\mathbf{R}^6$ 中的向量，所以 Col $\boldsymbol{B}$ 的维数不可能超过 6，即 $\text{rank}\,\boldsymbol{B}$ 不能超过 6. ■

下面这道例题为子空间的直观化提供了很好的方法. 第 6 章中我们将要学到，Row $\boldsymbol{A}$ 和 Nul $\boldsymbol{A}$ 的公共向量只有零向量，两个空间相互“垂直”. 这个事实对Row $\boldsymbol{A}^{\mathrm{T}}$ (= Col $\boldsymbol{A}$) 和 Nul $\boldsymbol{A}^{\mathrm{T}}$ 也成立. 例题 4 中的图 4-24 则给出了一般情形下的直观图像. (习题 29 中阐明了同时研究 $\boldsymbol{A}$ 和 $\boldsymbol{A}^{\mathrm{T}}$ 的意义.)

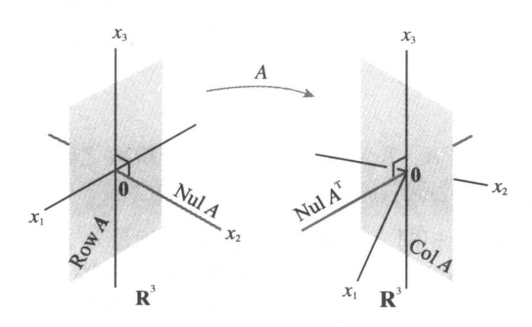


图 4-24 矩阵 A 确定的子空间

【例题 4】 设 $\boldsymbol{A}=\begin{bmatrix}3 & 0 & -1\\ 3 & 0 & -1\\ 4 & 0 & 5\end{bmatrix}$，容易验证 Nul $\boldsymbol{A}$ 是 x_2 - 轴，Row $\boldsymbol{A}$ 是 x_1x_3 平面，Col $\boldsymbol{A}$是满足方程 $x_1-x_2=0$ 的平面，并且 Nul $\boldsymbol{A}^{\mathrm{T}}$是$(1,-1,0)$的全体倍数的集合. 图 4-24 表明 Nul $\boldsymbol{A}$ 和 Row $\boldsymbol{A}$ 包含在线性变换 $\boldsymbol{x}\mapsto \boldsymbol{A}\boldsymbol{x}$ 的定义域中；该映射的像 Col $\boldsymbol{A}$ 与 Nul $\boldsymbol{A}^{\mathrm{T}}$ 一起包含在 $\mathbf{R}^3$的另一个副本中. ■

4.6.3 在方程组中的应用

秩定理是处理线性方程组相关信息的一个强有力的工具. 现实生活中一些实际问题需要运用线性方程的知识，但在表述上不一定明确采用线性代数的术语，比如矩阵、子空间和维数等等. 下面的例题模仿这种陈述方式描述了一个实际问题.

266

【例题 5】 某科学家求出了一个含有 40 个方程和 42 个变量的齐次方程组的两个解. 这两个解之间没有倍数关系，并且其他任意的解可以用这两个解的适当倍数之和来构造. 那么该科学家是否可以断定对应的(即,具有相同系数的)非齐次方程组有解？

解： 可以. 设 $\boldsymbol{A}$ 是方程组的 40×42 的系数矩阵，题设表明这两个解是线性无关的并且可以张成 Nul $\boldsymbol{A}$，因此 $\dim \text{Nul}\,\boldsymbol{A}=2$. 又由秩定理可知，$\dim \text{Col}\,\boldsymbol{A}=42-2=40$.

因为 $\mathbf{R}^{40}$ 是 $\mathbf{R}^{40}$ 中唯一的维数为 40 的子空间，因此 Col $\boldsymbol{A}$ 就是 $\mathbf{R}^{40}$. 这就意味着每个非齐次方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解. ■

4.6.4 秩和可逆矩阵定理

有了与矩阵相关的各种向量空间的概念，我们可以为可逆矩阵定理添加新的命题. 沿用 2.3 节中可逆矩阵定理中原有的编号，这些新命题列举如下：

【定理】 可逆矩阵定理(续)

设 $\boldsymbol{A}$ 是 $n\times n$ 矩阵，则下列的命题均等价于命题 $\boldsymbol{A}$ 是可逆矩阵：

(m) $\boldsymbol{A}$ 的列构成 $\mathbf{R}^n$ 的一组基；

(n) Col $\boldsymbol{A}=\mathbf{R}^n$；

(o) dim Col $\boldsymbol{A}=n$；

(p) rank $\boldsymbol{A}=n$；

(q) Nul $\boldsymbol{A}=\{\mathbf{0}\}$；

(r) dim Nul $\boldsymbol{A}=0$.

267

证明： 命题(m)等价于(e)和(h)，后两者分别关于线性无关和张成. 其余五个命题与定理前面的命题有下面的蕴含关系链：

$$(g)\Rightarrow(n)\Rightarrow(o)\Rightarrow(p)\Rightarrow(r)\Rightarrow(q)\Rightarrow(d)$$

命题(g)说的是方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 对任意 $\boldsymbol{b}\in\mathbf{R}^n$ 都至少有一个解，它蕴含(n)，因为 Col $\boldsymbol{A}$ 恰好是使方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解的全体 $\boldsymbol{b}$ 所构成的集合. (n)⇒(o)⇒(p)的蕴含关系可以由维数和秩的定义得到. 如果 $\boldsymbol{A}$ 的秩等于列数，那么根据秩定理有 dim Nul $\boldsymbol{A}=\mathbf{0}$，所以 Nul $\boldsymbol{A}=\{\mathbf{0}\}$. 于是(p)⇒(r)⇒(q). 此外，由(q)可知方程 $\boldsymbol{Ax}=\mathbf{0}$ 仅有平凡解，即命题(d). 因为已知(d)和(g)都等价于 $\boldsymbol{A}$ 可逆这一命题，定理证毕. ■

由于矩阵 $\boldsymbol{A}$ 的行空间即 $\boldsymbol{A}^{\mathrm{T}}$ 的列空间，所以关于 $\boldsymbol{A}$ 的行空间的事实是显然的，因而没有列入可逆矩阵定理. 回顾可逆矩阵定理的命题(l)：$\boldsymbol{A}$ 可逆当且仅当 $\boldsymbol{A}^{\mathrm{T}}$ 可逆. 因此可逆矩阵定理中的每个命题也可以用 $\boldsymbol{A}^{\mathrm{T}}$ 陈述. 不过，那样的话，定理的长度将翻倍，得到 30 多个命题！

注记

本书讨论的许多算法有助于读者理解概念，也适合进行简单的手算. 不过，对于实际生活中大规模的问题，这些算法往往不可取.

秩的判定就是一个很好的例子. 将矩阵化简成阶梯形式，然后数出主元列，看起来非常容易，但除非在元素明确指定的矩阵上施行精确的算法，否则行运算可能会改变矩阵的显秩. 例如，如果矩阵 $\begin{bmatrix}5 & 7\\ 5 & x\end{bmatrix}$ 中的 x 没有被当作 7 精确地存储在计算机中，则该矩阵的秩可能是 1 也可能是 2，这取决于计算机是否把 $x-7$ 当成零.

在实际应用中，矩阵 $\boldsymbol{A}$ 的有效秩经常由 $\boldsymbol{A}$ 的奇异值分解来确定，我们将在 7.4 节中加以讨论. 奇异值分解能够可靠地求得 Col $\boldsymbol{A}$，Row $\boldsymbol{A}$，Nul $\boldsymbol{A}$ 和 Nul $\boldsymbol{A}^{\mathrm{T}}$ 空间中的基.

基础练习

下列矩阵是行等价的：

$$A = \begin{bmatrix} 2 & -1 & 1 & -6 & 8 \\ 1 & -2 & -4 & 3 & -2 \\ -7 & 8 & 10 & 3 & -10 \\ 4 & -5 & -7 & 0 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -2 & -4 & 3 & -2 \\ 0 & 3 & 9 & -12 & 12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

268

1. 求 rank A 和 dim Nul A.
2. 求 Col A 和 Row A 的基.
3. 为求 Nul A 的基，下一步该做什么？
4. 在 A^T 的行阶梯形矩阵中有多少个主元列？

习题 4.6

在习题 1 ~4 中，假设矩阵 A 行等价于 B. 不用计算，试列出 rank A 和 dim Nul A，然后求 Col A，Row A 和 Nul A 的基.

1. $A = \begin{bmatrix} 1 & -4 & 9 & -7 \\ -1 & 2 & -4 & 1 \\ 5 & -6 & 10 & 7 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & -1 & 5 \\ 0 & -2 & 5 & -6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

2. $A = \begin{bmatrix} 1 & -3 & 4 & -1 & 9 \\ -2 & 6 & -6 & -1 & -10 \\ -3 & 9 & -6 & -6 & -3 \\ 3 & -9 & 4 & 9 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -3 & 0 & 5 & -7 \\ 0 & 0 & 2 & -3 & 8 \\ 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

3. $A = \begin{bmatrix} 2 & -3 & 6 & 2 & 5 \\ -2 & 3 & -3 & -3 & -4 \\ 4 & -6 & 9 & 5 & 9 \\ -2 & 3 & 3 & -4 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -3 & 6 & 2 & 5 \\ 0 & 0 & 3 & -1 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & 1 & -3 & 7 & 9 & -9 \\ 1 & 2 & -4 & 10 & 13 & -12 \\ 1 & -1 & -1 & 1 & 1 & -3 \\ 1 & -3 & 1 & -5 & -7 & 3 \\ 1 & -2 & 0 & 0 & -5 & -4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & -3 & 7 & 9 & -9 \\ 0 & 1 & -1 & 3 & 4 & -3 \\ 0 & 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

5. 如果 3×8 矩阵 A 的秩为 3，试求 dim Nul A，dim Row A 和 rank A^T.
6. 如果 6×3 矩阵 A 的秩为 3，试求 dim Nul A，dim Row A 和 rank A^T.
7. 假设 4×7 矩阵 A 有 4 个主元列. 是否有 Col $A = \mathbb{R}^4$？是否有 Nul $A = \mathbb{R}^3$？说明理由.
8. 假设 5×6 矩阵 A 有四个主元列. dim Nul A 是多少？是否有 Col $A = \mathbb{R}^4$？为什么？
9. 如果一个 5×6 矩阵 A 的零空间是 4 维的，则 A 的列空间的维数是多少？
10. 如果一个 7×6 矩阵 A 的零空间是 5 维的，则 A 的列空间的维数是多少？
11. 如果一个 8×5 矩阵 A 的零空间是 2 维的，则 A 的行空间的维数是多少？
12. 如果一个 5×6 矩阵 A 的零空间是 4 维的，则 A 的行空间的维数是多少？
13. 如果 A 是 7×5 矩阵，则 A 的秩最大可能是多少？如果 A 是 5×7 矩阵，则 A 的秩最大可能是多少？解释你的答案.
14. 如果 A 是 4×3 矩阵，则 A 的行空间的维数最大可能是多少？如果 A 是 3×4 矩阵，则 A 的行空间的维数最大可能是多少？解释你的答案.
15. 如果 A 是 6×8 矩阵，则 Nul A 的维数最小可能是多少？
16. 如果 A 是 6×4 矩阵，则 Nul A 的维数最小可能是多少？

习题 17 和 18 中 A 是 $m \times n$ 矩阵，试判断下列各命题的真假，并说明理由.

17. a. $\boldsymbol{A}$ 的行空间与 $\boldsymbol{A}^{\mathrm{T}}$的列空间相同.
    b. 如果 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的任一阶梯形式，并且 $\boldsymbol{B}$ 有 3 个非零行，则 $\boldsymbol{A}$ 的前 3 行构成 Row $\boldsymbol{A}$ 的一组基.
    c. $\boldsymbol{A}$ 的行空间和列空间的维数相同，即使 $\boldsymbol{A}$ 不是方阵.
    d. $\boldsymbol{A}$ 的行空间的维数与零空间的维数的和等于 $\boldsymbol{A}$ 中的行数.
    e. 在计算机上施行行变换可能改变矩阵的显秩.

269

18. a. 如果 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的任一阶梯形式，则 $\boldsymbol{B}$ 的主元列构成 $\boldsymbol{A}$ 的列空间的一组基.
    b. 行变换保持 $\boldsymbol{A}$ 的行之间的线性相关关系.
    c. $\boldsymbol{A}$ 的零空间的维数等于 $\boldsymbol{A}$ 中非主元列的个数.
    d. $\boldsymbol{A}^{\mathrm{T}}$ 的行空间与 $\boldsymbol{A}$ 的列空间相同.
    e. 如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是行等价的，则它们的行空间相同.
19. 假设一个齐次方程组含有 5 个方程 6 个未知量，它的所有的解均是某个非零解的倍数，问如果任意选取方程右端的常数，该方程组是否一定有解？并加以解释.
20. 假设一个含有 6 个线性方程，8 个未知量的非齐次方程组有一个解含有 2 个自由变量，是否有可能通过改变方程右端的常数，使得新方程组不相容？并加以解释.
21. 假设一个非齐次方程组含有 9 个方程 10 个未知量，并且不论方程右端常数取值如何，该方程组均有解. 问对于与其相对应的齐次方程组，是否可以找到 2 个解并且它们之间没有倍数关系？试讨论.
22. 一个含有 10 个方程 12 个未知量的齐次方程组，它的所有的解是否有可能是某个固定解的倍数？试讨论.
23. 假设一个齐次方程组含有 12 个线性方程 8 个未知量，它有 2 个相互之间没有倍数关系的固定解，并且其他所有的解均是这 2 个解的线性组合. 问所有解的集合是否可以用少于 12 个方程的齐次线性方程组来描述？如果可以，需要多少？试讨论.
24. 一个含有 7 个方程 6 个未知量的非齐次方程组，是否有可能对于等号右端某组常数有唯一解？是否有可能对等号右端任意一组常数都有唯一解？试加以解释.
25. 某科研人员在解一个含 10 个方程 12 个未知量的非齐次线性方程组时，发现它有 3 个未知量是自由变量. 他是否可以断定，如果改变方程右端，它仍将有解？试讨论.
26. 在统计学理论中，一般要求矩阵为满秩的，也就是说，它的秩要尽可能地大. 试解释为什么一个行数多于列数的 $m\times n$ 矩阵是满秩的当且仅当它的列是线性无关的.

习题 27 ~ 29 涉及的是 $m\times n$ 矩阵 $\boldsymbol{A}$，以及由它确定的基本子空间.

27. 子空间 Row $\boldsymbol{A}$，Col $\boldsymbol{A}$，Nul $\boldsymbol{A}$，Row $\boldsymbol{A}^{\mathrm{T}}$，Col $\boldsymbol{A}^{\mathrm{T}}$ 和 Nul $\boldsymbol{A}^{\mathrm{T}}$，哪些在 $\mathbf{R}^m$ 中？哪些在 $\mathbf{R}^n$ 中？其中互不相同的子空间有多少个？
28. 验证下列等式：
    a. $\dim \operatorname{Row} \boldsymbol{A} + \dim \operatorname{Nul} \boldsymbol{A} = n$（$\boldsymbol{A}$ 的列数）
    b. $\dim \operatorname{Col} \boldsymbol{A} + \dim \operatorname{Nul} \boldsymbol{A}^{\mathrm{T}} = m$（$\boldsymbol{A}$ 的行数）
29. 利用习题 28 来解释为什么对任意 $\boldsymbol{b}\in\mathbf{R}^m$，方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解当且仅当方程 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{0}$ 只有平凡解？
30. 假设 $\boldsymbol{A}$ 是 $m\times n$ 阵，$\boldsymbol{b}\in\mathbf{R}^m$. 为使方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 相容，$\operatorname{rank}[\boldsymbol{A}\quad \boldsymbol{b}]$ 和 rank $\boldsymbol{A}$ 之间应该满足什么关系？

秩为 1 的矩阵在一些计算机算法以及理论研究中十分重要，比如第 7 章的奇异值分解. 可以证明，$m\times n$ 矩阵 $\boldsymbol{A}$ 的秩为 1 当且仅当它是一个外积，即 $\boldsymbol{A}=\boldsymbol{uv}^{\mathrm{T}}$，其中 $\boldsymbol{u}\in\mathbf{R}^m$，$\boldsymbol{v}\in\mathbf{R}^n$. 习题 31 ~ 33 解释了其中的缘由.

31. 验证 $\operatorname{rank} \boldsymbol{uv}^{\mathrm{T}}\leqslant 1$，其中 $\boldsymbol{u}=\begin{bmatrix}2\\-3\\5\end{bmatrix}, \boldsymbol{v}=\begin{bmatrix}a\\b\\c\end{bmatrix}$.

32. 设 $u=\begin{bmatrix}1\\2\end{bmatrix}$. 求 $v\in\mathbf{R}^3$，使 $\begin{bmatrix}1&-3&4\\2&-6&8\end{bmatrix}=uv^{\mathrm{T}}$.
33. 设 A 是满足 rank $A=1$ 的任意 2×3 矩阵，设 u 是 A 的第一列，假设 $u\neq\mathbf{0}$. 解释为什么在 $\mathbf{R}^3$ 中存在一个向量 v 使得 $A=uv^{\mathrm{T}}$？如果 A 的第一列为零，该构造应该如何修改？
34. 设 A 是一个秩大于零的 $m\times n$ 矩阵，并且 U 是 A 的阶梯形式. 解释为什么存在一个可逆矩阵 E 使得 $A=EU$，并且利用该事实把 A 写成秩为 1 的矩阵的和的形式. [提示：见2.4节定理10.]
35. [M]设 $A=\begin{bmatrix}7&-9&-4&5&3&-3&-7\\-4&6&7&-2&-6&-5&5\\5&-7&-6&5&-6&2&8\\-3&5&8&-1&-7&-4&8\\6&-8&-5&4&4&9&3\end{bmatrix}$.
   a. 构造矩阵 C 和 N，使得它们的列分别是 Col A 和 Nul A 的基，并且构造矩阵 R，使它的行构成 Row A 的一组基.

270

   b. 构造矩阵 M 使得它的列构成 Nul A^{T} 的一组基. 构造 $S=[R^{\mathrm{T}}\quad N]$ 和 $T=[C\quad M]$，解释为什么 S 和 T 是方阵，并验证 S 和 T 均是可逆的.
36. [M]对于一个秩至多为4的 6×7 随机整数矩阵 A，重做习题35. 产生 A 的一种方式就是构造一个 6×4 随机整数矩阵 J 和一个 4×7 随机整数矩阵 K，令 $A=JK$.（见本章末尾的补充题12以及学习指南中的矩阵生成程序.）
37. [M]设 A 是习题3和5中的矩阵. 构造矩阵 C，使得它的列是 A 的主元列，构造矩阵 R，使得它的行是 A 化成阶梯形式后的非零行. 计算 CR，并讨论你的发现.
38. [M]对秩分别为5、4和3的3个 5×7 随机整数矩阵重做习题37，猜测 CR 和 A 之间的关系，并证明你的猜想.

基础练习答案

1. A 有2个主元列，所以 rank $A=2$. 由于 A 共有5列，dim Nul $A=5-2=3$.
2. A 的主元列是前两列，所以 Col A 的一组基是

$$\{a_1,a_2\}=\left\{\begin{bmatrix}2\\1\\-7\\4\end{bmatrix},\begin{bmatrix}-1\\-2\\8\\-5\end{bmatrix}\right\}$$

   B 的非零行构成 Row A 的一组基，即 $\{(1,-2,-4,3,-2),(0,3,9,-12,12)\}$. 在这个特定的例子中，恰巧 A 的任意2行均构成行空间的基，这是因为该行空间是2维的，并且 A 的任意行不是其他行的倍数. 一般而言，可以用 A 的阶梯形矩阵的非零行来构成 Row A 的一组基，而不是 A 本身的行.
3. 对于 Nul A，下一步是继续通过对 B 进行行化简，从而得到 A 的一个阶梯形式.
4. 由秩定理可知，rank $A^{\mathrm{T}}=$ rank A，因为 Col $A^{\mathrm{T}}=$ Row A，所以 A^{T} 有2个主元位置.

4.7　基 变 换

选定 n 维向量空间 V 的一组基 $\mathcal{B}$，则与之对应的从 V 到 $\mathbf{R}^n$ 上的坐标映射就确定

了V上的一个坐标系，V中的每个$\boldsymbol{x}$均可由它的$\mathcal{B}$-坐标向量$[\boldsymbol{x}]_{\mathcal{B}}$唯一确定.[1]

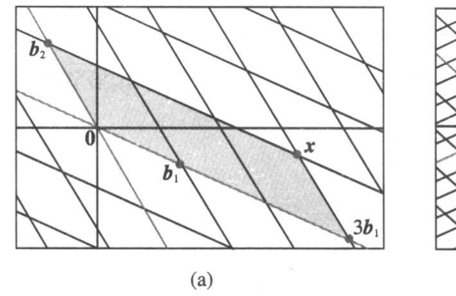
在某些应用中，问题最初可能用基$\mathcal{B}$来描述，但要解答它却需要将$\mathcal{B}$转化成新的基$\mathcal{C}$(在第5章和第7章会给出这样的例子). 这样每个向量就指派了一个新的$\mathcal{C}$-坐标向量. 本节中，我们将探讨对于V中的$\boldsymbol{x}$，$[\boldsymbol{x}]_{\mathcal{B}}$和$[\boldsymbol{x}]_{\mathcal{C}}$之间有何联系.

271

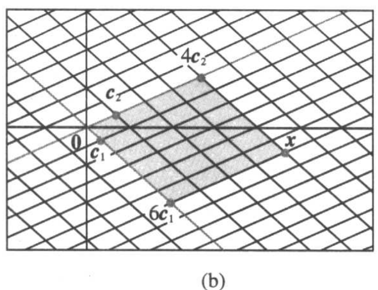
为了使这个问题更直观，考虑图4-25中的两个坐标系. 在图4-25(a)中，$\boldsymbol{x}=3\boldsymbol{b}_1+\boldsymbol{b}_2$，但在图4-25(b)中，同一个$\boldsymbol{x}$满足$\boldsymbol{x}=6\boldsymbol{c}_1+4\boldsymbol{c}_2$，即

$$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}3\\1\end{bmatrix}\text{且}[\boldsymbol{x}]_{\mathcal{C}}=\begin{bmatrix}6\\4\end{bmatrix}$$

我们的目标是找出两个坐标向量之间的联系. 下面的例题1告诉我们如何做到这一点，前提是我们知道$\boldsymbol{b}_1$和$\boldsymbol{b}_2$是怎样由$\boldsymbol{c}_1$和$\boldsymbol{c}_2$构造的.



(a)



(b)

图4-25 同一向量空间的两个坐标系

【例题1】 假设$\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$和$\mathcal{C}=\{\boldsymbol{c}_1,\boldsymbol{c}_2\}$是向量空间$V$的两组基，满足

$$\boldsymbol{b}_1=4\boldsymbol{c}_1+\boldsymbol{c}_2\quad\text{且}\quad\boldsymbol{b}_2=-6\boldsymbol{c}_1+\boldsymbol{c}_2\tag{1}$$

如果

$$\boldsymbol{x}=3\boldsymbol{b}_1+\boldsymbol{b}_2\tag{2}$$

即$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix}3\\1\end{bmatrix}$，求$[\boldsymbol{x}]_{\mathcal{C}}$.

解： 将$\mathcal{C}$确定的坐标映射作用在(2)中的$\boldsymbol{x}$上. 由于坐标映射是线性变换，我们有

$$[\boldsymbol{x}]_{\mathcal{C}}=[3\boldsymbol{b}_1+\boldsymbol{b}_2]_{\mathcal{C}}=3[\boldsymbol{b}_1]_{\mathcal{C}}+[\boldsymbol{b}_2]_{\mathcal{C}}$$

我们可以将这个向量方程写成一个矩阵方程，把线性组合中的向量作为矩阵的列：

$$[\boldsymbol{x}]_{\mathcal{C}}=[\,[\boldsymbol{b}_1]_{\mathcal{C}}\quad[\boldsymbol{b}_2]_{\mathcal{C}}]\begin{bmatrix}3\\1\end{bmatrix}\tag{3}$$

一旦我们知道矩阵的列向量，就可以利用上式求得$[\boldsymbol{x}]_{\mathcal{C}}$. 由(1)有

$$[\boldsymbol{b}_1]_{\mathcal{C}}=\begin{bmatrix}4\\1\end{bmatrix}\quad\text{且}\quad[\boldsymbol{b}_2]_{\mathcal{C}}=\begin{bmatrix}-6\\1\end{bmatrix}$$

1. 可以把$[\boldsymbol{x}]_{\mathcal{B}}$看作是$\boldsymbol{x}$的“名字”，它列出了用$\mathcal{B}$中基向量的线性组合表示$\boldsymbol{x}$时所用的权重.

272 于是，(3)式给出了答案：

$$[\boldsymbol{x}]_{\mathcal{C}}=\begin{bmatrix}4 & -6\\ 1 & 1\end{bmatrix}\begin{bmatrix}3\\ 1\end{bmatrix}=\begin{bmatrix}6\\ 4\end{bmatrix}$$

$\boldsymbol{x}$ 的这个 $\mathcal{C}$－坐标与图 4-25 中 $\boldsymbol{x}$ 的坐标一致. ■

公式(3)的推导可以加以推广并得到下面的结论(见习题 15 和 16).

【定理 15】　设 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$和 $\mathcal{C}=\{\boldsymbol{c}_1,\cdots,\boldsymbol{c}_n\}$是向量空间 V 的两组基，则存在唯一的 $n\times n$ 矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 使得

$$[\boldsymbol{x}]_{\mathcal{C}}=\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\boldsymbol{x}]_{\mathcal{B}} \tag{4}$$

其中 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的列是 $\mathcal{B}$ 中向量的 $\mathcal{C}$－坐标向量，即

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}=[[\boldsymbol{b}_1]_{\mathcal{C}}\quad[\boldsymbol{b}_2]_{\mathcal{C}}\quad\cdots\quad[\boldsymbol{b}_n]_{\mathcal{C}}] \tag{5}$$

定理 15 中的矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 称为 **$\mathcal{B}$ 到 $\mathcal{C}$ 的坐标变换矩阵**(change-of-coordinates matrix from $\mathcal{B}$ to $\mathcal{C}$). 向量的 $\mathcal{B}$－坐标乘以 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 就得到对应的 $\mathcal{C}$－坐标.[1]图 4-26 解释了坐标变换方程(4).

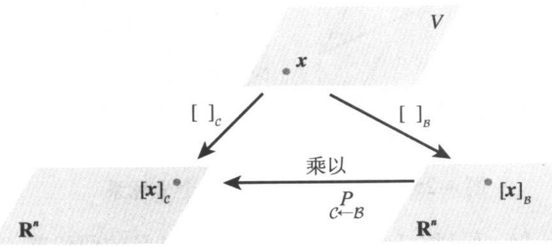


图 4-26　V 的两个坐标系

矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的列向量是线性无关的，这是因为它们是线性无关集 $\mathcal{B}$ 的坐标向量(见 4.4 节习题 25). 又由于 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 是方阵，根据可逆矩阵定理，$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 一定可逆. 方程(4)的两端同时左乘$(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}$得到

273

$$(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}[\boldsymbol{x}]_{\mathcal{C}}=[\boldsymbol{x}]_{\mathcal{B}}$$

因此矩阵$(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}$将 $\mathcal{C}$－坐标转换成 $\mathcal{B}$－坐标，即

$$(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}=\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} \tag{6}$$

$\mathbf{R}^n$的基变换

设 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$，且 ε 是 $\mathbf{R}^n$ 中的标准基$\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$，则$[\boldsymbol{b}_1]_{\varepsilon}=\boldsymbol{b}_1$，$\mathcal{B}$ 中其他向量也类似. 此时，$\underset{\varepsilon\leftarrow\mathcal{B}}{P}$ 与 4.4 节中介绍的坐标变换矩阵 $P_{\mathcal{B}}$相同，即

$$P_{\mathcal{B}}=[\boldsymbol{b}_1\quad\boldsymbol{b}_2\quad\cdots\quad\boldsymbol{b}_n]$$

---

1. 为了记住这个矩阵的构造，可以将 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\boldsymbol{x}]_{\mathcal{B}}$ 看成 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 中列的线性组合. 这个矩阵－向量积是一个 $\mathcal{C}$－坐标向量，因此 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的列也应该是 $\mathcal{C}$－坐标向量.

为了在 $\mathbf{R}^n$ 中的两个非标准坐标基之间进行坐标变换，我们需要用到定理15. 该定理说明，为了求解基变换问题，需要求出原来基关于新的基的坐标向量.

【例题2】 设 $\boldsymbol{b}_1=\begin{bmatrix}-9\\1\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-5\\-1\end{bmatrix},\boldsymbol{c}_1=\begin{bmatrix}1\\-4\end{bmatrix},\boldsymbol{c}_2=\begin{bmatrix}3\\-5\end{bmatrix}$，考虑 $\mathbf{R}^n$ 的基 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 和 $\mathcal{C}=\{\boldsymbol{c}_1,\boldsymbol{c}_2\}$，求 $\mathcal{B}$ 到 $\mathcal{C}$ 的坐标变换矩阵.

解： 矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 包含 $\boldsymbol{b}_1$ 和 $\boldsymbol{b}_2$ 的 $\mathcal{C}$-坐标向量. 设 $[\boldsymbol{b}_1]_{\mathcal{C}}=\begin{bmatrix}x_1\\x_2\end{bmatrix},[\boldsymbol{b}_2]_{\mathcal{C}}=\begin{bmatrix}y_1\\y_2\end{bmatrix}$. 则由定义知

$$[\boldsymbol{c}_1\quad\boldsymbol{c}_2]\begin{bmatrix}x_1\\x_2\end{bmatrix}=\boldsymbol{b}_1\quad 且\quad[\boldsymbol{c}_1\quad\boldsymbol{c}_2]\begin{bmatrix}y_1\\y_2\end{bmatrix}=\boldsymbol{b}_2$$

为了同时解出两个方程组，将 $\boldsymbol{b}_1$ 和 $\boldsymbol{b}_2$ 添加到系数矩阵中，并进行行化简：

$$[\boldsymbol{c}_1\quad\boldsymbol{c}_2\ \vdots\ \boldsymbol{b}_1\quad\boldsymbol{b}_2]=\left[\begin{array}{cc:cc}1&3&-9&-5\\-4&-5&1&-1\end{array}\right]\sim\left[\begin{array}{cc:cc}1&0&6&4\\0&1&-5&-3\end{array}\right]\tag{7}$$

因此，

$$[\boldsymbol{b}_1]_{\mathcal{C}}=\begin{bmatrix}6\\-5\end{bmatrix}\quad 且\quad[\boldsymbol{b}_2]_{\mathcal{C}}=\begin{bmatrix}4\\-3\end{bmatrix}$$

所求的坐标变换矩阵是

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}=[[\boldsymbol{b}_1]_{\mathcal{C}}\quad[\boldsymbol{b}_2]_{\mathcal{C}}]=\begin{bmatrix}6&4\\-5&-3\end{bmatrix}$$ ■

经观察可发现，例题2所求的 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 在(7)式中已给出. 不过这并不奇怪，因为 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的第一列是通过将 $[\boldsymbol{c}_1\quad\boldsymbol{c}_2\ \vdots\ \boldsymbol{b}_1]$ 行化简成 $[I\ \vdots\ [\boldsymbol{b}_1]_{\mathcal{C}}]$ 而得到的，第2列也类似. 因此

$$\boxed{[\boldsymbol{c}_1\quad\boldsymbol{c}_2\ \vdots\ \boldsymbol{b}_1\quad\boldsymbol{b}_2]\sim[I\ \vdots\ \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}]}$$

274

$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的求法可以类推到 $\mathbf{R}^n$，用来求 $\mathbf{R}^n$ 中任意两组基之间的坐标变换矩阵.

【例题3】 设 $\boldsymbol{b}_1=\begin{bmatrix}1\\-3\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-2\\4\end{bmatrix},\boldsymbol{c}_1=\begin{bmatrix}-7\\9\end{bmatrix},\boldsymbol{c}_2=\begin{bmatrix}-5\\7\end{bmatrix}$，考虑 $\mathbf{R}^2$ 的基 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 和 $\mathcal{C}=\{\boldsymbol{c}_1,\boldsymbol{c}_2\}$.

a. 求 $\mathcal{C}$ 到 $\mathcal{B}$ 的坐标变换矩阵；

b. 求 $\mathcal{B}$ 到 $\mathcal{C}$ 的坐标变换矩阵.

解：

a. 注意所求的是 $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$ 而不是 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$，计算

$$[\boldsymbol{b}_1\quad\boldsymbol{b}_2\ \vdots\ \boldsymbol{c}_1\quad\boldsymbol{c}_2]=\left[\begin{array}{cc:cc}1&-2&-7&-5\\-3&4&9&7\end{array}\right]\sim\left[\begin{array}{cc:cc}1&0&5&3\\0&1&6&4\end{array}\right]$$

所以有

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}=\begin{bmatrix}5&3\\6&4\end{bmatrix}$$

b. 根据(a)和性质(6)(将 $\mathcal{B}$ 和 $\mathcal{C}$ 互换)，

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = (\underset{\mathcal{B}\leftarrow\mathcal{C}}{P})^{-1} = \frac{1}{2}\begin{bmatrix} 4 & -3 \\ -6 & 5 \end{bmatrix} = \begin{bmatrix} 2 & -3/2 \\ -3 & 5/2 \end{bmatrix}$$ ■

我们还可以利用 $P_{\mathcal{B}}$ 和 $P_{\mathcal{C}}$ 来描述坐标变换矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$，这里 $P_{\mathcal{B}}$ 和 $P_{\mathcal{C}}$ 分别是将 $\mathcal{B}$-坐标和 $\mathcal{C}$-坐标变换到标准坐标的坐标变换矩阵. 回忆对 $\mathbf{R}^n$中的任意 x 都有：

$$P_{\mathcal{B}}[x]_{\mathcal{B}} = x,\quad P_{\mathcal{C}}[x]_{\mathcal{C}} = x,\quad 且\quad [x]_{\mathcal{C}} = P_{\mathcal{C}}^{-1}x$$

因此

$$[x]_{\mathcal{C}} = P_{\mathcal{C}}^{-1}x = P_{\mathcal{C}}^{-1}P_{\mathcal{B}}[x]_{\mathcal{B}}$$

属于 $\mathbf{R}^n$，坐标变换矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 可以通过计算 $P_{\mathcal{C}}^{-1}P_{\mathcal{B}}$ 得到. 实际上，对于比 2×2 更大的矩阵，使用类似于例题 3 的算法比先求 $P_{\mathcal{C}}^{-1}$ 再计算 $P_{\mathcal{C}}^{-1}P_{\mathcal{B}}$更快. 见 2.2 节习题 12.

基础练习

1. 设 $\mathcal{F}=\{f_1,f_2\}$和 $\mathcal{G}=\{g_1,g_2\}$是向量空间 V 的两组基，矩阵 P 的列向量为$[f_1]_{\mathcal{G}}$和$[f_2]_{\mathcal{G}}$. 对于 V 中全体 v，P 能满足下列哪一个方程？

   (i) $[v]_{\mathcal{F}} = P[v]_{\mathcal{G}}$　　(ii) $[v]_{\mathcal{G}} = P[v]_{\mathcal{F}}$

275

2. 设 $\mathcal{B}$ 和 $\mathcal{C}$ 同例题 1，利用该例题的结果求从 $\mathcal{C}$ 到 $\mathcal{B}$ 的坐标变换矩阵.

习题 4.7

1. 设 $\mathcal{B}=\{b_1,b_2\}$和 $\mathcal{C}=\{c_1,c_2\}$是向量空间 V 的两组基，并且假设 $b_1=6c_1-2c_2, b_2=9c_1-4c_2$.
   a. 求从 $\mathcal{B}$ 到 $\mathcal{C}$ 的坐标变换矩阵.
   b. 设 $x=-3b_1+2b_2$，利用(a)求$[x]_{\mathcal{C}}$.
2. 设 $\mathcal{B}=\{b_1,b_2\}$和 $\mathcal{C}=\{c_1,c_2\}$是向量空间 V 的两组基，并且假设 $b_1=-c_1+4c_2, b_2=5c_1-3c_2$.
   a. 求从 $\mathcal{B}$ 到 $\mathcal{C}$ 的坐标变换矩阵.
   b. 对于 $x=5b_1+3b_2$，求$[x]_{\mathcal{C}}$.
3. 设 $\mathcal{U}=\{u_1,u_2\}$和 $\mathcal{W}=\{w_1,w_2\}$是 V 的基，并且设矩阵 P 的列向量是$[u_1]_{\mathcal{W}}$和$[u_2]_{\mathcal{W}}$，对于 V 中全体 v，下列哪个方程能够被 P 满足？

   (i) $[x]_{\mathcal{U}} = P[x]_{\mathcal{W}}$　　(ii) $[x]_{\mathcal{W}} = P[x]_{\mathcal{U}}$
4. 设 $\mathcal{A}=\{a_1,a_2,a_3\}$和 $\mathcal{D}=\{d_1,d_2,d_3\}$是 V 的基，并且假设 $P=[\,[d_1]_{\mathcal{A}}\ \ [d_2]_{\mathcal{A}}\ \ [d_3]_{\mathcal{A}}]$. 对于 V 中所有的 x，下列哪个方程能够被 P 满足？

   (i) $[x]_{\mathcal{A}} = P[x]_{\mathcal{D}}$　　(ii) $[x]_{\mathcal{D}} = P[x]_{\mathcal{A}}$
5. 设 $\mathcal{A}=\{a_1,a_2,a_3\}$和 $\mathcal{B}=\{b_1,b_2,b_3\}$是向量空间 V 的基，并且假设 $a_1=4b_1-b_2, a_2=-b_1+b_2+b_3, a_3=b_2-2b_3$.
   a. 求从 $\mathcal{A}$ 到 $\mathcal{B}$ 的坐标变换矩阵.
   b. 设 $x=3a_1+4a_2+a_3$，求$[x]_{\mathcal{B}}$.
6. 设 $\mathcal{D}=\{d_1,d_2,d_3\}$和 $\mathcal{F}=\{f_1,f_2,f_3\}$是向量空间 V 的基，并且假设 $f_1=2d_1-d_2+d_3, f_2=3d_2+d_3, f_3=-3d_1+2d_3$.
   a. 求从 $\mathcal{F}$ 到 $\mathcal{D}$ 的坐标变换矩阵.
   b. 对于 $x=f_1-2f_2+2f_3$，求$[x]_{\mathcal{D}}$.

在习题 7~10 中，设 $\mathcal{B}=\{b_1,\ b_2\}$和 $\mathcal{C}=\{c_1,\ c_2\}$是 $\mathbf{R}^2$的两组基，对于每道习题，分别求从 $\mathcal{B}$ 到 $\mathcal{C}$ 和从 $\mathcal{C}$ 到 $\mathcal{B}$ 的坐标变换矩阵.

7. $\boldsymbol{b}_1=\begin{bmatrix}7\\5\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-3\\-1\end{bmatrix},\boldsymbol{c}_1=\begin{bmatrix}1\\-5\end{bmatrix},\boldsymbol{c}_2=\begin{bmatrix}-2\\2\end{bmatrix}$ 8. $\boldsymbol{b}_1=\begin{bmatrix}-1\\8\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}1\\-5\end{bmatrix},\boldsymbol{c}_1=\begin{bmatrix}1\\4\end{bmatrix},\boldsymbol{c}_2=\begin{bmatrix}1\\1\end{bmatrix}$

9. $\boldsymbol{b}_1=\begin{bmatrix}-6\\-1\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}2\\0\end{bmatrix},\boldsymbol{c}_1=\begin{bmatrix}2\\-1\end{bmatrix},\boldsymbol{c}_2=\begin{bmatrix}6\\-2\end{bmatrix}$ 10. $\boldsymbol{b}_1=\begin{bmatrix}7\\-2\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}2\\-1\end{bmatrix},\boldsymbol{c}_1=\begin{bmatrix}4\\1\end{bmatrix},\boldsymbol{c}_2=\begin{bmatrix}5\\2\end{bmatrix}$

在习题 11 和 12 中，$\mathcal{B}$ 和 $\mathcal{C}$ 是向量空间 V 的基，判断各命题的真假，并说明理由.

11. a. 坐标变换矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的列向量是 $\mathcal{C}$ 中向量的 $\mathcal{B}$ - 坐标向量.
    b. 如果 $V=\mathbb{R}^n$，$\mathcal{C}$ 是 V 的标准基，则 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 与 4.4 节中介绍的坐标变换矩阵 $P_{\mathcal{B}}$ 相同.
12. a. $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的列向量是线性无关的.
    b. 如果 $V=\mathbb{R}^2,\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 且 $\mathcal{C}=\{\boldsymbol{c}_1,\boldsymbol{c}_2\}$，则将 $[\boldsymbol{c}_1\quad\boldsymbol{c}_2\quad\boldsymbol{b}_1\quad\boldsymbol{b}_2]$ 行化简成 $[I\quad P]$ 所产生的矩阵 P 满足：对任意 $\boldsymbol{x}\in V$，都有 $[\boldsymbol{x}]_{\mathcal{B}}=P[\boldsymbol{x}]_{\mathcal{C}}$.
13. 在 $\mathbb{P}_2$ 中，求从基 $\mathcal{B}=\{1-2t+t^2,3-5t+4t^2,2t+3t^2\}$ 到标准基 $\mathcal{C}=\{1,t,t^2\}$ 的坐标变换矩阵，然后求 $-1+2t$ 的 $\mathcal{B}$ - 坐标向量.
14. 在 $\mathbb{P}_2$ 中，求从基 $\mathcal{B}=\{1-3t^2,2+t-5t^2,1+2t\}$ 到标准基的坐标变换矩阵，然后将 t^2 写成 $\mathcal{B}$ 中多项式的线性组合的形式.

习题 15 和 16 给出了定理 15 的一个证明，补充每一步证明的理由.

15. 已知 $\boldsymbol{v}\in V$，存在数量 $x_1,\cdots,x_n$ 满足
$$\boldsymbol{v}=x_1\boldsymbol{b}_1+x_2\boldsymbol{b}_2+\cdots+x_n\boldsymbol{b}_n$$
这是因为(a)______. 对 $\boldsymbol{v}$ 作用基 $\mathcal{C}$ 所确定的坐标映射，可以得到
$[\boldsymbol{v}]_{\mathcal{C}}=x_1[\boldsymbol{b}_1]_{\mathcal{C}}+x_2[\boldsymbol{b}_2]_{\mathcal{C}}+\cdots+x_n[\boldsymbol{b}_n]_{\mathcal{C}}$
这是因为(b)______. 该方程可以写成
$$[\boldsymbol{v}]_{\mathcal{C}}=[[\boldsymbol{b}_1]_{\mathcal{C}}\quad[\boldsymbol{b}_2]_{\mathcal{C}}\quad\cdots\quad[\boldsymbol{b}_n]_{\mathcal{C}}]\begin{bmatrix}x_1\\\vdots\\x_n\end{bmatrix}\tag{8}$$
这是根据定义(c)______. (8)表明(5)中的矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 满足：对任意 $\boldsymbol{v}\in V$ 都有 $[\boldsymbol{v}]_{\mathcal{C}}=\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\boldsymbol{v}]_{\mathcal{B}}$，这是因为(8)式右端的向量就是(d)______.
16. 假设 Q 是满足下列条件的任意矩阵
$$\text{对任意 }\boldsymbol{v}\in V\text{，有 }[\boldsymbol{v}]_{\mathcal{C}}=Q[\boldsymbol{v}]_{\mathcal{B}}\tag{9}$$
在(9)中令 $\boldsymbol{v}=\boldsymbol{b}_1$，则(9)表明 $[\boldsymbol{b}_1]_{\mathcal{C}}$ 是 Q 的第一列，这是因为(a)______. 类似地，对于 $k=2,\cdots,n$，Q 的第 k 列是(b)______，这是因为(c)______. 这就表明在定理 15 中由(5)定义的矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 是唯一满足条件(4)的矩阵.

276

17. [M]设 $\mathcal{B}=\{\boldsymbol{x}_0,\cdots,\boldsymbol{x}_6\},\mathcal{C}=\{\boldsymbol{y}_0,\cdots\boldsymbol{y}_6\}$，其中 $\boldsymbol{x}_k$ 是函数 $\cos^k t$，$\boldsymbol{y}_k$ 是函数 $\cos kt$. 4.5 节习题 34 表明 $\mathcal{B}$ 和 $\mathcal{C}$ 都是向量空间 $H=\text{Span}\{\boldsymbol{x}_0,\cdots,\boldsymbol{x}_6\}$ 的基.
    a. 令 $P=[[\boldsymbol{y}_0]_{\mathcal{B}}\cdots[\boldsymbol{y}_6]_{\mathcal{B}}]$，计算 P^{-1}.
    b. 解释 P^{-1} 的列向量为什么是 $\boldsymbol{x}_0,\cdots,\boldsymbol{x}_6$ 的 $\mathcal{C}$ - 坐标向量. 然后，利用这个坐标向量写出 $\cos t$ 幂的三角恒等表达式，其中只允许用 $\mathcal{C}$ 中的函数.
18. [M](需要微积分知识)[1] 回忆微积分的基本知识，积分

1. 习题 17,18 以及前几节中 5 道相关习题的思想来自 Auburn 大学 Jack W. Rogers, Jr. 的一篇论文，此文发表于 1995 年 8 月的国际线性代数学会会议. 见“Applications of Linear Algebra in Calculus”, *American Mathematical Monthly* **104** (1), 1997.

$$\int(5\cos^3 t-6\cos^4 t+5\cos^5 t-12\cos^6 t)\,dt \tag{10}$$

计算起来相当繁琐(通常的方法是重复利用分部积分法和半角公式). 利用习题17中矩阵 $\boldsymbol{P}$ 或 $\boldsymbol{P}^{-1}$ 对(10)做变换, 然后计算积分.

19. [M]设

$$\boldsymbol{P}=\begin{bmatrix}1&2&-1\\-3&-5&0\\4&6&1\end{bmatrix},\ \boldsymbol{v}_1=\begin{bmatrix}-2\\2\\3\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}-8\\5\\2\end{bmatrix},\boldsymbol{v}_3=\begin{bmatrix}-7\\2\\6\end{bmatrix}$$

a. 求 $\mathbf{R}^3$ 的一组基 $\{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3\}$, 使得 $\boldsymbol{P}$ 是从 $\{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3\}$ 到基 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 的坐标变换矩阵. [提示: $\underset{\mathcal{C}\leftarrow\mathcal{B}}{\boldsymbol{P}}$ 的列向量代表什么?]

b. 求 $\mathbf{R}^3$ 的一组基 $\{\boldsymbol{w}_1,\boldsymbol{w}_2,\boldsymbol{w}_3\}$, 使得 $\boldsymbol{P}$ 是从 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 到基 $\{\boldsymbol{w}_1,\boldsymbol{w}_2,\boldsymbol{w}_3\}$ 的坐标变换矩阵.

20. 设 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\},\mathcal{C}=\{\boldsymbol{c}_1,\boldsymbol{c}_2\}$ 和 $\mathcal{D}=\{\boldsymbol{d}_1,\boldsymbol{d}_2\}$ 是一个2维向量空间的基.

a. 写出 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{\boldsymbol{P}}$, $\underset{\mathcal{D}\leftarrow\mathcal{C}}{\boldsymbol{P}}$, $\underset{\mathcal{D}\leftarrow\mathcal{B}}{\boldsymbol{P}}$ 之间的关系式, 并验证你的结果.

b. [M]在 $\mathbf{R}^2$ 中选取三组不同的基(见习题7~10.), 利用矩阵软件给出(a)中的关系式, 或者验证你的结果.

基础练习答案

1. 因为 $\boldsymbol{P}$ 的列向量是 $\mathcal{G}$-坐标向量, 形如 $\boldsymbol{Px}$ 的向量一定是 $\mathcal{G}$-坐标向量, 因此 $\boldsymbol{P}$ 满足方程(ii).
2. 例题1中求出的坐标向量表明

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{\boldsymbol{P}}=[\,[\boldsymbol{b}_1]_\mathcal{C}\quad[\boldsymbol{b}_2]_\mathcal{C}]=\begin{bmatrix}4&-6\\1&1\end{bmatrix}$$

因此

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{\boldsymbol{P}}=(\underset{\mathcal{C}\leftarrow\mathcal{B}}{\boldsymbol{P}})^{-1}=\frac{1}{10}\begin{bmatrix}1&6\\-1&4\end{bmatrix}=\begin{bmatrix}0.1&0.6\\-0.1&0.4\end{bmatrix}$$

4.8 在差分方程中的应用

随着高性能计算机的广泛应用, 越来越多的科技、工程问题基于离散或者数字化的数据, 而非连续数据来实现. 差分方程往往是分析这类数据的得力工具. 即便在我们用微分方程为某个连续处理过程建模时, 其数值解也经常由一个相关的差分方程而产生.

277

本节着重强调了线性差分方程的几个基本性质, 它们可以用线性代数的知识加以解释.

4.8.1 离散时间信号

离散时间信号的向量空间 $\mathbf{S}$ 已经在4.1节有所介绍. $\mathbf{S}$ 中的一个**信号**(signal)就是一个定义在整数上的函数, 它可以看作是一个数列, 比如 $\{y_k\}$. 图4-27给出了三个典型的信号, 它们的通项分别是 $(0.7)^k$, 1^k 和 $(-1)^k$.

数字信号常见于电子学和控制系统工程学中, 而在生物学、物理学、经济学、人口统计学以及其他众多领域里, 只要在离散的时间区间对某个过程进行测量或采样, 也都会得到离散数据序列. 如果指定一个过程的开始时间, 把信号记成 $(y_0,y_1,y_2,\cdots)$ 的形式往往更方便, 对于 $k<0$, 可以假定项 y_k 是零或者干脆省略.

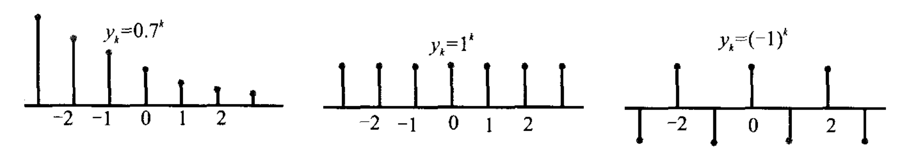


图 4-27 S 中的三个信号

【例题 1】 CD 播放机所播放的声音十分清晰，它是以每秒 44 100 次的采样频率采样得到的音乐，见图 4-28. 每采样一次，音乐信号的振幅都被记录成一个数值 y_k，原始的音乐由许多不同频率的声音组成，然而，序列$\{y_k\}$所包含的信息可以重现频率在 20 000 赫兹以下的所有声音，而高出这个频率的声音人耳根本听不到. ■

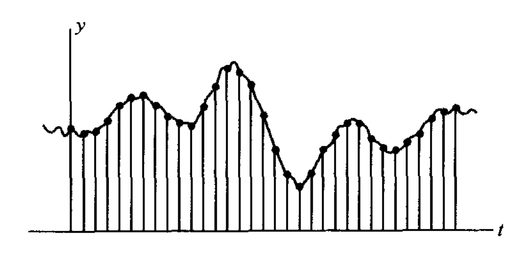


图 4-28 音乐信号的采样数据

278

4.8.2 信号空间 S 中的线性无关性

为简化记号，我们考虑 S 中一个只含三个信号$\{u_k\}$,$\{v_k\}$和$\{w_k\}$的集合. 仅当方程

$$c_1u_k + c_2v_k + c_3w_k = 0 \qquad \text{对任意的 } k \tag{1}$$

蕴含 $c_1 = c_2 = c_3 = 0$ 时，这三个信号线性无关. “对任意的 k” 是指对任意的整数，包括正数、负数和零. 也可以考虑从 $k=0$ 开始的信号，此时，“对任意的 k” 是指对任意的整数 $k \geqslant 0$.

假设 c_1，c_2，c_3满足(1)，则(1)中的方程对任意 3 个连续的 k 值都成立，比如说，k，$k+1$，$k+2$. 因此，(1)蕴含

$$c_1u_{k+1} + c_2v_{k+1} + c_3w_{k+1} = 0 \qquad \text{对任意的 } k$$

和

$$c_1u_{k+2} + c_2v_{k+2} + c_3w_{k+2} = 0 \qquad \text{对任意的 } k$$

因此 c_1，c_2，c_3满足

$$\begin{bmatrix} u_k & v_k & w_k \\ u_{k+1} & v_{k+1} & w_{k+1} \\ u_{k+2} & v_{k+2} & w_{k+2} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \text{对任意的 } k \tag{2}$$

该方程组的系数矩阵称为信号的卡索拉蒂**矩阵**(Casorati matrix)，该矩阵的行列式称为$\{u_k\}$,$\{v_k\}$和$\{w_k\}$的卡索拉蒂行列式(Casoratian determinant). 如果至少存在一个 k 值使卡索拉蒂矩阵可逆，则(2)将蕴含 $c_1 = c_2 = c_3 = 0$，这就证明了这三个信号线性无关.

【例题2】　证明 1^k，$(-2)^k$ 和 3^k 是线性无关的信号.

解： 卡索拉蒂矩阵是

$$\begin{bmatrix} 1^k & (-2)^k & 3^k \\ 1^{k+1} & (-2)^{k+1} & 3^{k+1} \\ 1^{k+2} & (-2)^{k+2} & 3^{k+2} \end{bmatrix}$$

利用行变换很容易证明该矩阵总是可逆的. 但是更快捷的方法是用某个具体的值代替 k，比如 $k=0$，然后对所得的数值矩阵进行行化简：

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 3 \\ 1 & 4 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 0 & 3 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 0 & 0 & 10 \end{bmatrix}$$

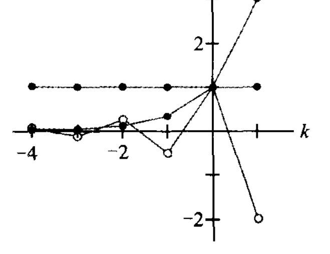


图 4-29　信号 1^k，$(-2)^k$ 和 3^k

对 $k=0$，卡索拉蒂矩阵可逆. 因此 1^k，$(-2)^k$ 和 3^k 是线性无关的. ■

如果卡索拉蒂矩阵不可逆，则待检验的信号可能线性相关，也可能线性无关(见习题33). 但是可以证明，如果这些信号是同一个齐次差分方程(见下文)的所有解，那么，或者对任意的 k，卡索拉蒂矩阵都可逆并且信号线性无关，或者对任意的 k，卡索拉蒂矩阵都不可逆并且信号线性相关. 对于这个事实，学习指南中利用线性变换给出了一个漂亮的证明.

279

4.8.3　线性差分方程

给定数量 $a_0,\cdots,a_n$，其中 a_0,a_n 非零，以及信号 $\{z_k\}$，称方程

$$a_0y_{k+n}+a_1y_{k+n-1}+\cdots+a_{n-1}y_{k+1}+a_ny_k=z_k \quad \text{对任意的 } k \tag{3}$$

为 **n 阶线性差分方程**(linear difference equation of order n)或**线性递推关系**(linear recurrence relation of order n). 如果 $\{z_k\}$ 是零序列，该方程是**齐次的**(homogeneous)，否则就是**非齐次的**(nonhomogeneous).

【例题3】　在数字信号处理中，形如(3)的差分方程描述了一个**线性滤波器**(linear filter)，其中的 $a_0,\cdots,a_n$ 称作**滤波器系数**(filter coefficient). 如果把 $\{y_k\}$ 和 $\{z_k\}$ 分别看成是输入和输出，则对应的齐次方程的解就是被过滤且转化为零信号的信号. 我们将两个不同的信号输入滤波器

$$0.35y_{k+2}+0.5y_{k+1}+0.35y_k=z_k$$

其中 0.35 是 $\sqrt{2}/4$ 的近似值. 第一个信号是由连续信号 $y=\cos(\pi t/4)$ 在 t 取整数值时采样而得到，见图 4-30a. 该离散信号就是

$$\{y_k\}=\{\cdots,\cos(0),\cos(\pi/4),\cos(2\pi/4),\cos(3\pi/4),\cdots\}$$

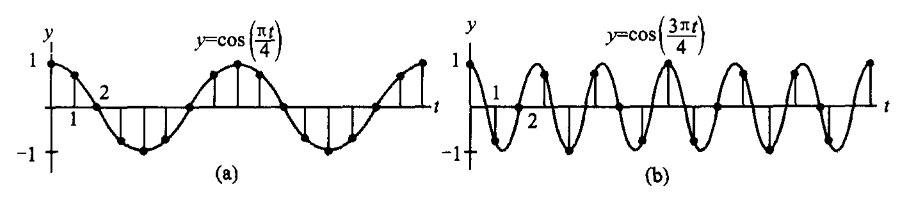


图 4-30　不同频率的离散信号

简单起见，以 ±0.7 替换 $\pm\sqrt{2}/2$，得到

$$\{y_k\}=\{\cdots,1,0.7,0,-0.7,-1,-0.7,0,0.7,1,0.7,0,\cdots\}$$
$$\uparrow k=0$$

表 4-2 给出了输出序列$\{z_k\}$的计算过程，其中 0.35(0.7)是$(\sqrt{2}/4)(\sqrt{2}/2)=0.25$ 的简写. 可以发现，输出序列$\{z_k\}$恰好是输入序列$\{y_k\}$后移一项所得的序列.

280

表 4-2 滤波器输出的计算过程

k	y_k	y_{k+1}	y_{k+2}	$0.35y_k \quad +0.5y_{k+1} \quad +0.35y_{k+2}=z_k$
0	1	0.7	0	0.35(1) +0.5(0.7) +0.35(0) =0.7
1	0.7	0	-0.7	0.35(0.7) +0.5(0) +0.35(-0.7) =0
2	0	-0.7	-1	0.35(0) +0.5(-0.7) +0.35(-1) = -0.7
3	-0.7	-1	-0.7	0.35(-0.7) +0.5(-1) +0.35(-0.7) = -1
4	-1	-0.7	0	0.35(-1) +0.5(-0.7) +0.35(0) = -0.7
5	-0.7	0	0.7	0.35(-0.7) +0.5(0) +0.35(0.7) =0
⋮	⋮			⋮

另一个输入信号由高频信号 $y=\cos(3\pi t/4)$ 产生，见图 4-30b. 使用与前一信号相同的采样频率，可以得到新的输入序列：

$$\{w_k\}=\{\cdots,1,-0.7,0,0.7,-1,0.7,0,-0.7,1,-0.7,0,\cdots\}$$
$$\uparrow k=0$$

将$\{w_k\}$作为滤波器的输入时，输出的就是零序列. 这样的滤波器称作**低通滤波器**(low-pass filter)，它允许$\{y_k\}$通过，却滤掉了高频的$\{w_k\}$. ■

在许多实际应用中，差分方程(3)右端的序列$\{z_k\}$通常是指定的，满足(3)的序列$\{y_k\}$称为方程的一个**解**(solution)，下面的例题说明了如何求齐次方程的解.

【例题 4】 齐次差分方程的解常常具有形式 $y_k=r^k$，其中 r 是某个数量. 试求下列方程的解：

$$y_{k+3}-2y_{k+2}-5y_{k+1}+6y_k=0 \qquad \text{对于任意的 } k \tag{4}$$

解：用 r^k替换方程中的 y_k，并且对方程左端进行因式分解：

$$r^{k+3}-2r^{k+2}-5r^{k+1}+6r^k=0 \tag{5}$$
$$r^k\ (r^3-2r^2-5r+6)\ =0$$
$$r^k\ (r-1)\ (r+2)\ (r-3)\ =0 \tag{6}$$

因为(5)等价于(6)，所以 r^k满足差分方程(4)当且仅当 r 满足(6). 因此 1^k，$(-2)^k$ 和 3^k都是(4)的解. 例如，为了验证 3^k是(4)的解，可以计算

$$3^{k+3}-2\cdot3^{k+2}-5\cdot3^{k+1}+6\cdot3^k$$
$$=3^k(27-18-15+6)=0 \quad \text{对于任意的 } k$$ ■

一般地，一个非零信号 r^k满足齐次差分方程

$$y_{k+n}+a_1y_{k+n-1}+\cdots+a_{n-1}y_{k+1}+a_ny_k=0 \qquad \text{对于任意的 } k$$

281

当且仅当 r 是下列**辅助方程**(auxiliary equation)的根：

$$r^n+a_1r^{n-1}+\cdots+a_{n-1}r+a_n=0$$

我们不考虑 r 是辅助方程的重根的情形. 如果辅助方程有复根，则差分方程有

$s^k\cos k\omega$ 和 $s^k\sin k\omega$ 形式的解，其中 s 和 ω 为常数，例题3就属于这种情形.

4.8.4 线性差分方程的解集

给定 $a_1,\cdots,a_n$，考虑由

$$w_k = y_{k+n} + a_1y_{k+n-1} + \cdots + a_{n-1}y_{k+1} + a_ny_k$$

确定的将信号 $\{y_k\}$ 映到 $\{w_k\}$ 的映射 T：$\mathbf{S}\to\mathbf{S}$. 容易验证 T 是一个线性变换，这就表明齐次方程

$$y_{k+n} + a_1y_{k+n-1} + \cdots + a_{n-1}y_{k+1} + a_ny_k = 0 \qquad \text{对于任意的 } k$$

的解集是 T 的核(被 T 映射成零信号的信号集合)，因此解集是 $\mathbf{S}$ 的一个子空间，解的任意线性组合仍然是一个解.

下面的定理是一个简单却基本的结论，它将引出关于差分方程的解集更多的信息.

【定理16】 如果 $a_n\neq0$，并且给定 $\{z_k\}$，只要指定 $y_0,\cdots,y_{n-1}$，则方程

$$y_{k+n} + a_1y_{k+n-1} + \cdots + a_{n-1}y_{k+1} + a_ny_k = z_k \qquad \text{对于任意的 } k \tag{7}$$

有唯一解.

证明： 如果 $y_0,\cdots,y_{n-1}$ 已经指定，利用(7)可定义

$$y_n = z_0 - [a_1y_{n-1} + \cdots + a_{n-1}y_1 + a_ny_0]$$

现在 $y_0,\cdots,y_n$ 已经指定，利用(7)可以继续定义 y_{n+1}. 一般地，对 $k\geqslant0$，我们可以利用递推关系

$$y_{n+k} = z_k - [a_1y_{k+n-1} + \cdots + a_ny_k] \tag{8}$$

来定义 y_{n+k}；对于 $k<0$，可以利用递推关系

$$y_k = \frac{1}{a_n}z_k - \frac{1}{a_n}[y_{k+n} + a_1y_{k+n-1} + \cdots + a_{n-1}y_{k+1}] \tag{9}$$

来定义 y_k. 这样就得到满足(7)的一个信号. 反之，如果一个信号对任意的 k 都满足(7)，则该信号必然满足(8)和(9)，所以(7)的解是唯一的. ■

【定理17】 n 阶齐次线性差分方程

$$y_{k+n} + a_1y_{k+n-1} + \cdots + a_{n-1}y_{k+1} + a_ny_k = 0 \qquad \text{对于任意的 } k \tag{10}$$

282

的全体解所构成的集合 H 是一个 n 维向量空间.

证明： 如前所述，H 是 $\mathbf{S}$ 的一个子空间，这是因为 H 是一个线性变换的核. 对于 H 中的 $\{y_k\}$，令 $F\{y_k\}$ 是 $\mathbf{R}^n$ 中的向量 $(y_0,y_1,\cdots,y_{n-1})$. 容易验证 F：$H\to\mathbf{R}^n$ 是线性变换. 给定 $\mathbf{R}^n$ 中的任意向量 $(y_0,y_1,\cdots,y_{n-1})$，定理16表明，H 中存在唯一的信号 $\{y_k\}$，满足 $F\{y_k\}=(y_0,y_1,\cdots,y_{n-1})$，这就意味着 F 是从 H 到 $\mathbf{R}^n$ 上的一对一的线性变换，即 F 是一个同构，因此 $\dim H=\dim\mathbf{R}^n=n$(见4.5节习题32). ■

【例题5】 求下列差分方程的解集的一组基：

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \qquad \text{对于任意的 } k$$

解： 我们所学的线性代数知识现在要发挥作用了！由例题2和4我们已经知道 1^k，$(-2)^k$ 和 3^k 是线性无关的解. 一般情况下，直接验证信号集合是否张成解空间非常困难，但在这里没有问题，因为有两个关键定理可以保证——那就是定理17，它说明该解空间是3维的，以及4.5节的基本定理，它说明 n 维空间中的含 n 个向量的线性无关集自然构成一组基. 所以 1^k，$(-2)^k$ 和 3^k 构成解空间的一组基. ■

表示(10)的"通解"的标准方法是给出全体解所构成的子空间的一组基，这组基通常称为(10)的**基础解系**(fundamental set of solutions). 实际上，如果你能找出满足(10)的 n 个线性无关的信号，按照例题 5 中的解释，它们能够自动张成 n 维的解空间.

4.8.5 非齐次方程

非齐次差分方程

$$y_{k+n}+a_1y_{k+n-1}+\cdots+a_{n-1}y_{k+1}+a_ny_k=z_k \qquad \text{对于任意的 } k \tag{11}$$

的通解可以写成(11)的一个特解加上对应齐次方程(10)的基础解系的任意线性组合. 这个事实类似于 1.5 节中结论：$\boldsymbol{Ax}=\boldsymbol{b}$ 和 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解集是平行的. 这两个结论有相同的解释：映射 $\boldsymbol{x}\to\boldsymbol{Ax}$ 是线性的，而方程(11)中将信号 $\{y_k\}$变换成 $\{z_k\}$的映射也是线性的，见习题 35.

【例题 6】 证明信号 $y_k=k^2$ 满足差分方程

$$y_{k+2}-4y_{k+1}+3y_k=-4k \qquad \text{对于任意的 } k \tag{12}$$

并且给出该方程全体解的一个刻画.

283

解：用 k^2代替方程(12)左端的 y_k：

$$(k+2)^2-4(k+1)^2+3k^2=(k^2+4k+4)-4(k^2+2k+1)+3k^2=-4k$$

所以 k^2确实是(12)的一个解. 接下来求解齐次方程

$$y_{k+2}-4y_{k+1}+3y_k=0 \tag{13}$$

辅助方程

$$r^2-4r+3=(r-1)(r-3)=0$$

的根是 $r=1$，3. 因此齐次差分方程的两个解是 1^k和 3^k. 显然它们相互没有倍数关系，因此是线性无关的信号. 根据定理 17，解空间是 2 维的，所以 1^k和 3^k构成(13)的解集的一组基. 再利用非齐次方程(12)的特解 k^2，我们可由$\{1^k,3^k\}$构造出(12)的通解：

$$k^2+c_11^k+c_23^k, \quad \text{或} \quad k^2+c_1+c_23^k$$

图 4-31 给出了这两个解集的几何表示. 图中的每个点都对应着 **S** 中的一个信号. ■

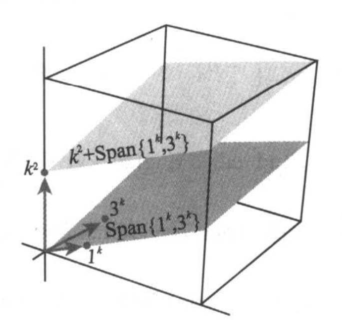


图 4-31 差分方程(12)和(13)的解集

4.8.6 化简为一阶方程组

研究齐次 n 阶线性差分方程的一种近代方法是用一个等价的一阶差分方程组来替代它，其形式为：

$$x_{k+1} = Ax_k \qquad \text{对于任意的 } k$$

其中向量 $x_k \in \mathbf{R}^n$，且 A 是一个 $n \times n$ 矩阵.

在1.10节中我们已经研究过(向量值)差分方程的一个简单例子，更多的例子将会在4.9节和5.6节涉及.

【例题7】 将下列差分方程写成一阶方程组：

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \qquad \text{对于任意的 } k$$

解：对于每个 k，令

$$x_k = \begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \end{bmatrix}$$

题设的差分方程表明 $y_{k+3} = -6y_k + 5y_{k+1} + 2y_{k+2}$，所以

$$x_{k+1} = \begin{bmatrix} y_{k+1} \\ y_{k+2} \\ y_{k+3} \end{bmatrix} = \begin{bmatrix} 0 + y_{k+1} + 0 \\ 0 + 0 + y_{k+2} \\ -6y_k + 5y_{k+1} + 2y_{k+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 5 & 2 \end{bmatrix} \begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \end{bmatrix}$$

284 即，

$$x_{k+1} = Ax_k \quad \text{对于任意的 } k\text{，其中} \quad A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 5 & 2 \end{bmatrix}$$

■

一般情况下，方程

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0 \qquad \text{对于任意的 } k$$

可以写成对任意 k 都有 $x_{k+1} = Ax_k$，其中

$$x_k = \begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+n-1} \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix}$$

更多阅读

Hamming, R. W., *Digital Filters*, 2nd ed. (Englewood Cliffs, NJ: Prentice-Hall, 1983), pp. 1-37.

Kelly, W. G., and A. C. Peterson, *Difference Equations*, 2nd ed. (San Diego: Harcourt-Academic Press, 2001).

Mickens, R. E., *Difference Equations*, 2nd ed. (New York: Van Nostrand Reinhold, 1990), pp. 88-141.

Oppenheim, A. V., and A. S. Willsky, *Signals and Systems*, 2nd ed, (Upper Saddle River, NJ: Prentice-Hall, 1997), pp. 1-14, 21-30, 38-43.

基础练习

可以证明信号 2^k，$3^k \sin \frac{k\pi}{2}$ 和 $3^k \cos \frac{k\pi}{2}$ 是方程

$$y_{k+3}-2y_{k+2}+9y_{k+1}-18y_k=0$$

的解. 证明这 3 个信号构成该差分方程解集的一组基.

习题 4.8

验证习题 1、2 中的信号是相应的差分方程的解.

1. $2^k, (-4)^k; y_{k+2}+2y_{k+1}-8y_k=0$
2. $3^k, (-3)^k; y_{k+2}-9y_k=0$

证明习题 3~6 中的信号构成相应的差分方程解的基.

3. 习题 1 中的信号和方程.
4. 习题 2 中的信号和方程.
5. $(-3)^k, k(-3)^k; y_{k+2}+6y_{k+1}+9y_k=0$

285

6. $5^k\cos\frac{k\pi}{2}, 5^k\sin\frac{k\pi}{2}; y_{k+2}+25y_k=0$

在习题 7~12 中, 假设列出的信号是给定的差分方程的解, 判断这些信号是否构成方程的解空间的一组基, 利用定理予以验证.

7. $1^k, 2^k, (-2)^k; y_{k+3}-y_{k+2}-4y_{k+1}+4y_k=0$
8. $2^k, 4^k, (-5)^k, y_{k+3}-y_{k+2}-22y_{k+1}+40y_k=0$
9. $1^k, 3^k\cos\frac{k\pi}{2}, 3^k\sin\frac{k\pi}{2}; y_{k+3}-y_{k+2}+9y_{k+1}-9y_k=0$
10. $(-1)^k, k(-1)^k, 5^k; y_{k+3}-3y_{k+2}-9y_{k+1}-5y_k=0$
11. $(-1)^k, 3^k, y_{k+3}+y_{k+2}-9y_{k+1}-9y_k=0$
12. $1^k, (-1)^k; y_{k+4}-2y_{k+2}+y_k=0$

在习题 13~16 中, 求差分方程的解空间的基, 并证明所求的解张成解集.

13. $y_{k+2}-y_{k+1}+\frac{2}{9}y_k=0$
14. $y_{k+2}-7y_{k+1}+12y_k=0$
15. $y_{k+2}-25y_k=0$
16. $16y_{k+2}+8y_{k+1}-3y_k=0$

习题 17 和 18 中讨论的是一个由差分方程

$$Y_{k+2}-a(1+b)Y_{k+1}+abY_k=1 \tag{14}$$

所描述的国民经济的简单模型, 其中 Y_k是在 k 年内的总的国民收入, a 是一个小于 1 的常数, 称为消费的边缘倾向, b 是一个正的修正常数, 它描述了每年个人投资的速度如何影响消费者的消费.[1]

17. 如果 $a=0.9$, $b=4/9$, 求(14)的通解. 随着 k 的增大, Y_k 将如何变化? [提示: 首先求 $Y_k=T$形式的一个特解, 其中 T 是常数, 称为国民收入的平衡水平.]
18. 如果 $a=0.9, b=0.5$, 求(14)的通解.

轻质悬梁由 N 个间距为10 英寸的支点所支撑, 500 磅的重物悬挂于悬梁尾端, 距第 1 个支点 10 英寸, 如图 4-32 所示. 设 y_k是第 k 个点的力矩, 因而 $y_1=500$ 英寸 - 磅. 假设悬梁在第 N 个支撑点是刚性的, 并且力矩为 0. 则中间点的力矩满足下列三力矩方程:

$$y_{k+2}+4y_{k+1}+y_k=0, 对于\ k=1,2,\cdots,N-2 \tag{15}$$

---

1. 例如, 见 *Discrete Dynamical Systems*, by James T. Sandefur, (Oxford: Clarendon Press, 1990), pp. 267~276. 最初的加速乘数模型是由经济学家萨缪尔森提出的.

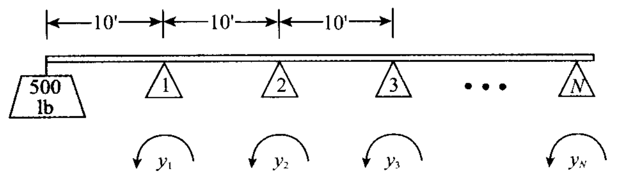


图 4-32　悬梁上的力矩

19. 求差分方程(15)的通解，并加以验证.
20. 求(15)的满足边界条件 $y_1=5000$ 和 $y_N=0$ 的特解(答案中包含 N).
21. 如果信号由对某个过程(化学反应、通过管道的热流和移动的机械臂等)的一系列测量而得到，它们通常包含了由测量误差产生的随机干扰. 降低这些干扰的一种标准预处理办法就是消除或过滤这些数据. 一个简单的过滤器是移动平均数，它利用 y_k 及其相邻两个值的平均数来替代 y_k：

$$\frac{1}{3}y_{k+1}+\frac{1}{3}y_k+\frac{1}{3}y_{k-1}=z_k \qquad 对于\ k=1,2,\cdots$$

假设一个信号 $y_k(k=0,\cdots,14)$ 是

9，5，7，3，2，4，6，5，7，6，8，10，9，5，7

利用这种过滤器计算 $z_1,\cdots,z_{13}$，并绘出原始信号和过滤后信号的图像.

22. 设 $\{y_k\}$ 是通过对连续信号 $2\cos\frac{\pi t}{4}+\cos\frac{3\pi t}{4}$ 在 $t=0$，1，2，…处取样得到的序列，如图 4-33所示. y_k 的值从 $k=0$ 开始分别是

3，0.7，0，-0.7，-3，-0.7，0，0.7，3，0.7，0，…

其中 0.7 是 $\sqrt{2}/2$ 的近似值.

a. 当 $\{y_k\}$ 输入到例题 3 中的过滤器时，计算出输出信号 $\{z_k\}$.

286

b. 请解释(a)的输出与例题 3 的计算结果有什么联系？为什么？

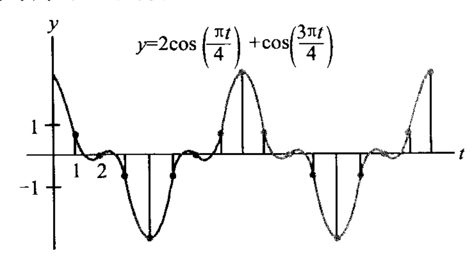


图 4-33　$2\cos\frac{\pi t}{4}+\cos\frac{3\pi t}{4}$ 的采样数据

习题 23 和 24 中讨论的是 $y_{k+1}-ay_k=b$ 形式的差分方程，其中 a，b 是适当的常数.

23. 设一份 \$10 000 贷款的利息是 1%/月，并且每月偿还 \$450. 贷款在 $k=0$ 月开始，第一次偿还在 1 个月后开始. 对 $k=0$，1，2，…，设 y_k 是第 k 个月偿还后贷款的未偿还部分的余额，则

$$\underset{新余额}{y_1}=\underset{应付余额}{10\ 000}+\underset{附加利息}{(0.01)10\ 000}-\underset{偿还}{450}$$

a. 写出 $\{y_k\}$ 满足的差分方程.

b. [M]创建关于 k 和在第 k 月的余额 y_k 的图表，并列出在创建图表的过程中你所用到的程序或键盘输入.

c. [M]最后一次偿还时，k 值是多少？最后一次将偿还多少？总共需要偿还多少？

24. 当 $k=0$ 时，将最初的投资 \$1000 存入一个活期账户中，利率为6%/年（每月的利率是0.005）．在最初的投资后的每个月再向该账户中注入 \$200．对于 $k=0, 1, 2, \cdots$，设 y_k 是在 k 月存入后账户内的总额.

 a. 写出一个 $\{y_k\}$ 满足的差分方程.

 b. [M]对于从 $k=0$ 到60，创建关于 k 和在第 k 月的余额 y_k 的图表，并列出在创建图表的过程中你所用到的程序或键盘输入.

 c. [M]两年（即24个月）后该账户内将会有多少存款？4年，5年呢？5年后的利息总额是多少？

在习题25～28中，证明给定的信号是差分方程的解，并求该差分方程通解.

25. $y_k=k^2; y_{k+2}+3y_{k+1}-4y_k=10k+7$

26. $y_k=1+k; y_{k+2}-8y_{k+1}+15y_k=8k+2$

27. $y_k=2-2k; y_{k+2}-\frac{9}{2}y_{k+1}+2y_k=3k+2$

28. $y_k=2k-4; y_{k+2}+\frac{3}{2}y_{k+1}-y_k=1+3k$

将习题29和30中的差分方程写成1阶方程组 $\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k$，对任意的 k

29. $y_{k+4}-6y_{k+3}+8y_{k+2}+6y_{k+1}-9y_k=0$

30. $y_{k+3}-\frac{3}{4}y_{k+2}+\frac{1}{16}y_k=0$

31. 下列差分方程是否是3阶的？并加以解释.

$$y_{k+3}+5y_{k+2}+6y_{k+1}=0$$

32. 下列差分方程是多少阶的？并解释你的答案.

$$y_{k+3}+a_1y_{k+2}+a_2y_{k+1}+a_3y_k=0$$

33. 设 $y_k=k^2, z_k=2k|k|$，信号 $\{y_k\}$ 和 $\{z_k\}$ 是否线性无关？对于 $k=0$，$k=-1$，$k=-2$，求出相应的卡索拉蒂矩阵 $C(k)$，并讨论你的结果.

34. 设 f，g，h 是定义在全体实数上的线性无关的函数，通过对函数

$$u_k=f(k), v_k=g(k), w_k=h(k)$$

在整数值处取样构造3个信号，这些信号在 $\mathbf{S}$ 中一定是线性无关的吗？并加以讨论.

35. 设 a，b 是非零数，证明：由 $T\{y_k\}=\{w_k\}$ 定义的映射是从 $\mathbf{S}$ 到 $\mathbf{S}$ 的线性变换，其中 $w_k=y_{k+2}+ay_{k+1}+by_k$.

36. 设 V 是个向量空间，$T: V\to V$ 是线性变换．给定 V 中的 $\boldsymbol{z}$，假设 V 中的 $\boldsymbol{x}_p$ 满足 $T(\boldsymbol{x}_p)=\boldsymbol{z}$，并且 $\boldsymbol{u}$ 是 T 的核中的任意一个向量，证明：$\boldsymbol{u}+\boldsymbol{x}_p$ 满足非齐次方程 $T(\boldsymbol{x})=\boldsymbol{z}$.

37. 设 $\mathbf{S}_0$ 是所有形如 $(y_0, y_1, y_2, \cdots)$ 的序列的向量空间，并且定义从 $\mathbf{S}_0$ 到 $\mathbf{S}_0$ 的线性变换 T 和 D 为

$$T(y_0, y_1, y_2, \cdots)=(y_1, y_2, y_3, \cdots)$$
$$D(y_0, y_1, y_2, \cdots)=(0, y_0, y_1, y_2, \cdots)$$

证明：$TD=I$（$\mathbf{S}_0$ 上的恒等变换），但 $DT\neq I$.

287

基础练习答案

检查卡索拉蒂矩阵

$$C(k)=\begin{bmatrix}2^k & 3^k\sin\frac{k\pi}{2} & 3^k\cos\frac{k\pi}{2}\\ 2^{k+1} & 3^{k+1}\sin\frac{(k+1)\pi}{2} & 3^{k+1}\cos\frac{(k+1)\pi}{2}\\ 2^{k+2} & 3^{k+2}\sin\frac{(k+2)\pi}{2} & 3^{k+2}\cos\frac{(k+2)\pi}{2}\end{bmatrix}$$

令 $k=0$，通过对矩阵进行行化简验证它有3个主元位置，因此是可逆的

$$C(0)=\begin{bmatrix}1 & 0 & 1\\ 2 & 3 & 0\\ 4 & 0 & -9\end{bmatrix}\sim\begin{bmatrix}1 & 0 & 1\\ 0 & 3 & -2\\ 0 & 0 & -13\end{bmatrix}$$

在 $k=0$ 处，卡索拉蒂矩阵是可逆的，所以这些信号是线性无关的. 由于有3个信号，并且差分方程的解空间 H 的维数是3(定理17)，由基本定理可知，这些信号构成 H 的一组基.

4.9　在马尔可夫链中的应用

本节介绍的马尔可夫链是生物、商业、化学、工程和物理等许多学科中应用广泛的一类数学模型. 这些模型常常用来描述以相同方式重复进行的实验或测量，每次实验的结果都在事先指定的几个可能结果之列，并且每次实验的结果都只依赖上一次实验.

例如，如果每年对某个城市及其郊区的人口进行测量，则向量

$$\boldsymbol{x}_0=\begin{bmatrix}0.60\\ 0.40\end{bmatrix}\tag{1}$$

就可以表明60%的人口住在城市，40%的人口住在郊区. $\boldsymbol{x}_0$中小数的和为1，代表该整个地区人口的总和. 这里我们使用百分率要比使用人口数更方便.

称一个只含非负分量并且各分量总和为1的向量为**概率向量**(probability vector)，一个**随机矩阵**(stochastic matrix)就是以概率向量为列向量的方阵. 一个**马尔可夫链**(Markov chain)就是一个概率向量序列 $\boldsymbol{x}_0,\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots$,以及随机矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{x}_1=\boldsymbol{P}\boldsymbol{x}_0,\boldsymbol{x}_2=\boldsymbol{P}\boldsymbol{x}_1,\boldsymbol{x}_3=\boldsymbol{P}\boldsymbol{x}_2,\cdots$$

因此马尔可夫链可以用下列1阶差分方程来描述：

$$\boldsymbol{x}_{k+1}=\boldsymbol{P}\boldsymbol{x}_k,\text{对于 }k=0,1,2,\cdots$$

如果 $\mathbf{R}^n$中向量的一个马尔可夫链描述了某个系统或者某个实验序列，则 $\boldsymbol{x}_k$中的分量依次列出了该系统处在全部 n 个可能状态上的概率，或者该实验结果是全部 n 个可能结果的概率，基于这一原因，$\boldsymbol{x}_k$也常称作**状态向量**(state vector).

288

【例题1】　在1.10节中，我们考察过一个城市和郊区之间的人口迁移模型，见图4-34. 每年城市和郊区之间的人口迁移受下列迁移矩阵 $\boldsymbol{M}$ 支配：

从：
城市　郊区　到：

$$\boldsymbol{M}=\begin{bmatrix}0.95 & 0.03\\ 0.05 & 0.97\end{bmatrix}\begin{matrix}\text{城市}\\ \text{郊区}\end{matrix}$$

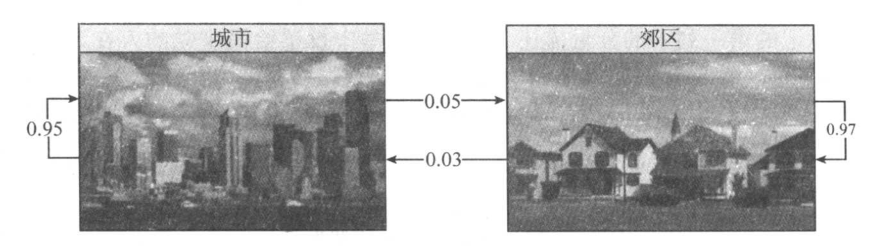


图 4-34 城市和郊区每年的迁移百分比

即每年有 5% 的城市人口迁往郊区，3% 的郊区人口迁往城市. M 的列向量是概率向量，因此 M 是随机矩阵. 假设 2000 年该地区的城市人口为 600 000，郊区人口为 400 000，则该地区人口的初始分布由上面(1)式中的 x_0 给出. 2001 年的人口分布是怎样的？2002 年呢？

解： 在 1.10 节例题 3 中，我们看到一年后人口向量 $\begin{bmatrix} 600\ 000 \\ 400\ 000 \end{bmatrix}$ 变化为

$$\begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix}\begin{bmatrix} 600\ 000 \\ 400\ 000 \end{bmatrix} = \begin{bmatrix} 582\ 000 \\ 418\ 000 \end{bmatrix}$$

如果将等式两端同时除以总人口 100 万，并且利用事实 $kMx = M(kx)$，我们就会发现

$$\begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix}\begin{bmatrix} 0.600 \\ 0.400 \end{bmatrix} = \begin{bmatrix} 0.582 \\ 0.418 \end{bmatrix}$$

向量 $x_1 = \begin{bmatrix} 0.582 \\ 0.418 \end{bmatrix}$ 给出了 2001 年的人口分布，即 58.2% 的人口在城市，41.8% 的人口在郊区. 类似地，2002 年的人口分布可以用向量 x_2 来描述，其中：

289

$$x_2 = Mx_1 = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix}\begin{bmatrix} 0.582 \\ 0.418 \end{bmatrix} = \begin{bmatrix} 0.565 \\ 0.435 \end{bmatrix}$$ ■

【例题 2】 假设在某选区内国会选举的投票结果由 $\mathbf{R}^3$ 中向量 x 给出：

$$x = \begin{bmatrix} \text{\% 民主党的选票 (D)} \\ \text{\% 共和党的选票 (R)} \\ \text{\% 自由者的选票 (L)} \end{bmatrix}$$

假设我们用这类向量来记录两年一次的国会选举的结果，并且每一次的选举结果只受前一次选举结果的影响，则描述每两年选举结果的向量序列可以作成一个马尔可夫链，例如，我们可以将对应的随机矩阵取成

$$\begin{array}{c} \text{从：} \\ \begin{array}{ccc} \text{D} & \text{R} & \text{L} \end{array} \quad \text{到：} \end{array}$$

$$P = \begin{bmatrix} 0.70 & 0.10 & 0.30 \\ 0.20 & 0.80 & 0.30 \\ 0.10 & 0.10 & 0.40 \end{bmatrix} \begin{array}{c} \text{D} \\ \text{R} \\ \text{L} \end{array}$$

“D”所标记的第一列中的分量描述了在某次选举中投票给民主党的人在下一次选举中的投票情况. 这里我们假设他们中的70%仍将投票给D，20%将投票给R，10%将投票给L. 对于$\boldsymbol{P}$中的其他列，也有类似的解释. 图4-35给出了该矩阵的图示.

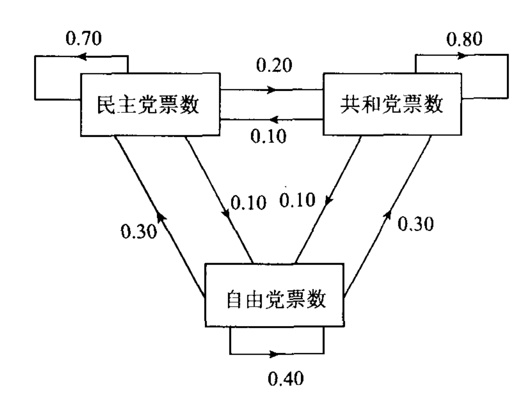


图4-35　从某次选举到下一次选举的选票变化

如果多年以来从一次选举到下次选举“转变”的百分比保持恒定，则给出选举结果的向量序列就构成一个马尔可夫链. 假设某次选举的结果是

290

$$\boldsymbol{x}_0=\begin{bmatrix}0.55\\0.40\\0.05\end{bmatrix}$$

估计下一次、再下一次选举可能的结果.

解：下一次、再下一次选举的结果可以分别用状态向量$\boldsymbol{x}_1$和$\boldsymbol{x}_2$来描述，其中

$$\boldsymbol{x}_1=P\boldsymbol{x}_0=\begin{bmatrix}0.70&0.10&0.30\\0.20&0.80&0.30\\0.10&0.10&0.40\end{bmatrix}\begin{bmatrix}0.55\\0.40\\0.05\end{bmatrix}=\begin{bmatrix}0.440\\0.445\\0.115\end{bmatrix}\quad\begin{matrix}44\%\text{将投票给 D}\\44.5\%\text{将投票给 R}\\11.5\%\text{将投票给 L}\end{matrix}$$

$$\boldsymbol{x}_2=P\boldsymbol{x}_1=\begin{bmatrix}0.70&0.10&0.30\\0.20&0.80&0.30\\0.10&0.10&0.40\end{bmatrix}\begin{bmatrix}0.440\\0.445\\0.115\end{bmatrix}=\begin{bmatrix}0.3870\\0.4785\\0.1345\end{bmatrix}\quad\begin{matrix}38.7\%\text{将投票给 D}\\47.8\%\text{将投票给 R}\\13.5\%\text{将投票给 L}\end{matrix}$$

为了理解为什么$\boldsymbol{x}_1$确实给出了下一次选举的结果，假设在“第一次”选举中有1000人投票，其中550人投票给D，400人投票给R，50人投票给L(见$\boldsymbol{x}_0$中的百分比). 在下一次选举中，550人中的70%会再次投票给D，400人中的10%将会把选票从R转投给D，50人中的30%将会把选票从L转投给D，因此D的总票数将会是

$$0.70(550)+0.10(400)+0.30(50)=385+40+15=440 \qquad (2)$$

因此下一次选举中将会有44%的选票投给D的候选人. (2)中的计算与$\boldsymbol{x}_1$中第1个分量的计算在本质上是相同的. 类似的计算可以求出$\boldsymbol{x}_1$的其他分量，以及$\boldsymbol{x}_2$的分量. ■

4.9.1　预测未来

马尔可夫链的一个最有趣的应用是研究链的长期行为. 例如，在例题2中多次

选举以后的情况会怎样(假设仍用给定的随机矩阵描述从一次选举到下一次选举的支持转变率)？或者，例题1中的人口分布“最后”究竟会怎样？在回答这些问题之前，我们先来看一个数值计算的例题.

【例题3】 设 $P=\begin{bmatrix}0.5 & 0.2 & 0.3\\0.3 & 0.8 & 0.3\\0.2 & 0 & 0.4\end{bmatrix}, x_0=\begin{bmatrix}1\\0\\0\end{bmatrix}$. 考虑一个由马尔可夫链 $x_{k+1}=Px_k$, $k=0,1,\cdots$描述的系统. 随着时间的变化，这个系统将如何变化？通过计算状态向量 $x_1,\cdots,x_{15}$来求解.

291

解：

$$x_1=Px_0=\begin{bmatrix}0.5 & 0.2 & 0.3\\0.3 & 0.8 & 0.3\\0.2 & 0 & 0.4\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix}=\begin{bmatrix}0.5\\0.3\\0.2\end{bmatrix}$$

$$x_2=Px_1=\begin{bmatrix}0.5 & 0.2 & 0.3\\0.3 & 0.8 & 0.3\\0.2 & 0 & 0.4\end{bmatrix}\begin{bmatrix}0.5\\0.3\\0.2\end{bmatrix}=\begin{bmatrix}0.37\\0.45\\0.18\end{bmatrix}$$

$$x_3=Px_2=\begin{bmatrix}0.5 & 0.2 & 0.3\\0.3 & 0.8 & 0.3\\0.2 & 0 & 0.4\end{bmatrix}\begin{bmatrix}0.37\\0.45\\0.18\end{bmatrix}=\begin{bmatrix}0.329\\0.525\\0.146\end{bmatrix}$$

进一步的计算结果如下所示，其中的分量四舍五入后保留4位或5位有效数字.

$$x_4=\begin{bmatrix}0.3133\\0.5625\\0.1242\end{bmatrix},\quad x_5=\begin{bmatrix}0.3064\\0.5813\\0.1123\end{bmatrix},\quad x_6=\begin{bmatrix}0.3032\\0.5906\\0.1062\end{bmatrix},\quad x_7=\begin{bmatrix}0.3016\\0.5953\\0.1031\end{bmatrix}$$

$$x_8=\begin{bmatrix}0.3008\\0.5977\\0.1016\end{bmatrix},\quad x_9=\begin{bmatrix}0.3004\\0.5988\\0.1008\end{bmatrix},\quad x_{10}=\begin{bmatrix}0.3002\\0.5994\\0.1004\end{bmatrix},\quad x_{11}=\begin{bmatrix}0.3001\\0.5997\\0.1002\end{bmatrix}$$

$$x_{12}=\begin{bmatrix}0.300\,05\\0.599\,85\\0.100\,10\end{bmatrix},x_{13}=\begin{bmatrix}0.300\,02\\0.599\,93\\0.100\,05\end{bmatrix},x_{14}=\begin{bmatrix}0.300\,01\\0.599\,96\\0.100\,02\end{bmatrix},x_{15}=\begin{bmatrix}0.300\,01\\0.599\,98\\0.100\,01\end{bmatrix}$$

这些向量看起来趋近于 $q=\begin{bmatrix}0.3\\0.6\\0.1\end{bmatrix}$. 当 k 从当前值变到下一个值时，上面所列的概率变化非常小. 注意，下面的计算是精确的(无舍入误差)：

$$Pq=\begin{bmatrix}0.5 & 0.2 & 0.3\\0.3 & 0.8 & 0.3\\0.2 & 0 & 0.4\end{bmatrix}\begin{bmatrix}0.3\\0.6\\0.1\end{bmatrix}=\begin{bmatrix}0.15+0.12+0.03\\0.09+0.48+0.03\\0.06+\quad 0+0.04\end{bmatrix}=\begin{bmatrix}0.30\\0.60\\0.10\end{bmatrix}=q$$

当系统处于状态 q 时，从一次测量到下一次测量无变化. ■

4.9.2　稳态向量

如果 P 是一个随机矩阵，则 P 的**稳态向量**(steady-state vector)或**平衡向量**(equilibrium vector)是满足下面条件的概率向量 q

$$Pq = q$$

292 可以证明每个随机矩阵都有稳态向量. 例题3中，q 就是 P 的一个稳态向量.

【例题4】　概率向量 $q=\begin{bmatrix}0.375\\0.625\end{bmatrix}$ 是例题1中的人口迁移矩阵 M 的一个稳态向量，这是因为

$$Mq=\begin{bmatrix}0.95 & 0.03\\0.05 & 0.97\end{bmatrix}\begin{bmatrix}0.375\\0.625\end{bmatrix}=\begin{bmatrix}0.35625+0.01875\\0.01875+0.60625\end{bmatrix}=\begin{bmatrix}0.375\\0.625\end{bmatrix}=q \qquad \blacksquare$$

如果例题1中的城市和郊区的总人口是100万，则例题4中的 q 表示375 000人在城市，625 000人在郊区. 一年之后，从城市迁出到郊区有(0.05)(375 000) = 18 750人，而从郊区迁入到城市有(0.03)(625 000) = 18 750人，结果，城市的人口将保持不变，同样，郊区的人口也不变.

下面的例题告诉我们如何求稳态向量.

【例题5】　设 $P=\begin{bmatrix}0.6 & 0.3\\0.4 & 0.7\end{bmatrix}$，求 P 的一个稳态向量.

解：首先，解方程 $Px=x$.

$$Px-x=0$$

$$Px-Ix=0 \qquad \text{回顾 1.4 节有：} Ix=x$$

$$(P-I)x=0$$

对于上述的 P，

$$P-I=\begin{bmatrix}0.6 & 0.3\\0.4 & 0.7\end{bmatrix}-\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}=\begin{bmatrix}-0.4 & 0.3\\0.4 & -0.3\end{bmatrix}$$

为求出 $(P-I)x=0$ 的全部解，行化简增广矩阵：

$$\begin{bmatrix}-0.4 & 0.3 & 0\\0.4 & -0.3 & 0\end{bmatrix}\sim\begin{bmatrix}-0.4 & 0.3 & 0\\0 & 0 & 0\end{bmatrix}\sim\begin{bmatrix}1 & -3/4 & 0\\0 & 0 & 0\end{bmatrix}$$

于是 $x_1=\frac{3}{4}x_2$，其中 x_2 是自由的. 通解是 $x_2\begin{bmatrix}3/4\\1\end{bmatrix}$.

下一步，为解空间选择一组简单的基. 显然可以选取 $\begin{bmatrix}3/4\\1\end{bmatrix}$，但一个更好(不含分数)的选择是 $w=\begin{bmatrix}3\\4\end{bmatrix}$ (对应于 $x_2=4$).

最后，在 $Px=x$ 的解集中找一个概率向量，这比较容易，因为每一个解都是 w

的倍数. 将 $\boldsymbol{w}$ 的每一分量除以各分量之和，可以得到

$$\boldsymbol{q}=\begin{bmatrix}3/7\\4/7\end{bmatrix}$$

验证一下，计算

293

$$\boldsymbol{Pq}=\begin{bmatrix}6/10 & 3/10\\4/10 & 7/10\end{bmatrix}\begin{bmatrix}3/7\\4/7\end{bmatrix}=\begin{bmatrix}18/70+12/70\\12/70+28/70\end{bmatrix}=\begin{bmatrix}30/70\\40/70\end{bmatrix}=\boldsymbol{q}$$ ■

下面的定理表明，例题 3 中的情况对于随机矩阵十分普遍. 我们称一个随机矩阵**正则**(regular)，如果矩阵的某个幂 $\boldsymbol{P}^k$ 只含有严格正的分量. 对于例题 3 中的 $\boldsymbol{P}$，我们有

$$\boldsymbol{P}^2=\begin{bmatrix}0.37 & 0.26 & 0.33\\0.45 & 0.70 & 0.45\\0.18 & 0.04 & 0.22\end{bmatrix}$$

因为 $\boldsymbol{P}^2$ 中的每个分量都是严格正的,所以 $\boldsymbol{P}$ 是一个正则的随机矩阵.

此外,我们称向量序列 $\{\boldsymbol{x}_k:k=1,2,\cdots\}$ 当 $k\to\infty$ 时**收敛**(converges)于向量 $\boldsymbol{q}$，如果对于选区充分大的 k，$\boldsymbol{x}_k$ 中的每个分量都能任意接近 $\boldsymbol{q}$ 中对应的分量.

【定理 18】 如果 $\boldsymbol{P}$ 是一个 $n\times n$ 正则随机向量，则 $\boldsymbol{P}$ 有唯一的稳态向量 $\boldsymbol{q}$. 而且，如果 $\boldsymbol{x}_0$ 是任意一个初始状态，并且有 $\boldsymbol{x}_{k+1}=\boldsymbol{Px}_k$，$k=0,1,2,\cdots$,则当 $k\to\infty$ 时马尔可夫链 $\{\boldsymbol{x}_k\}$ 收敛于 $\boldsymbol{q}$.

这个定理的证明可以参考任何一本关于马尔可夫链的标准教材. 定理的微妙之处在于初始状态不影响马尔可夫链的长期行为. 在后面(5.2 节)将会看到为什么对这里所讨论的几个随机矩阵这一事实都成立.

【例题 6】 在例题 2 中，假设选举结果构成一个马尔可夫链，从现在起许多年以后的选举中，共和党的候选人可能得到的选票的百分率是多少?

解：采用手算时，错误的方法是选定某个初始向量 $\boldsymbol{x}_0$，然后对于某个大的 k 值计算 $\boldsymbol{x}_0,\cdots,\boldsymbol{x}_k$，因为你不知道需要计算多少向量，也无法确定 $\boldsymbol{x}_k$ 中分量的极限值.

正确的方法是计算稳态向量，然后运用定理 18. 给定例题 2 中的 $\boldsymbol{P}$，从 $\boldsymbol{P}$ 的每个对角线元素减 1 得到 $\boldsymbol{P}-\boldsymbol{I}$，然后行化简增广矩阵:

$$[(\boldsymbol{P}-\boldsymbol{I})\quad \boldsymbol{0}]=\begin{bmatrix}-0.3 & 0.1 & 0.3 & 0\\0.2 & -0.2 & 0.3 & 0\\0.1 & 0.1 & -0.6 & 0\end{bmatrix}$$

294

回顾前面对小数的处理，将每一行乘以 10 可以简化计算.[1]

$$\begin{bmatrix}-3 & 1 & 3 & 0\\2 & -2 & 3 & 0\\1 & 1 & -6 & 0\end{bmatrix}\sim\begin{bmatrix}1 & 0 & -9/4 & 0\\0 & 1 & -15/4 & 0\\0 & 0 & 0 & 0\end{bmatrix}$$

---

1. **警告**：不要仅对矩阵 $\boldsymbol{P}$ 乘以 10，而应对方程 $(\boldsymbol{P}-\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$ 的增广矩阵乘以 10.

$(P-I)x=0$ 的通解是 $x_1=\frac{9}{4}x_3, x_2=\frac{15}{4}x_3$，其中 x_3是自由的. 取 $x_3=4$，我们得到解空间中一个各分量取整数的基，进而可以求出各分量和为 1 的稳态向量：

$$w=\begin{bmatrix}9\\15\\4\end{bmatrix},\quad q=\begin{bmatrix}9/28\\15/28\\4/28\end{bmatrix}\approx\begin{bmatrix}0.32\\0.54\\0.14\end{bmatrix}$$

q 的分量描述了多年以后选举中的选票分布(假设所给的随机矩阵始终描述从一次选举到下一次选举的变化). 因此，共和党候选人最终会得到 54% 的选票. ■

注记

你可能已经注意到如果 $x_{k+1}=Px_k, k=0,1,2,\cdots$，则

$$x_2=Px_1=P(Px_0)=P^2x_0$$

一般地有

$$x_k=P^kx_0 \qquad 对于 k=0,1,\cdots$$

为了计算出一个具体的向量，比如说 x_3，直接去计算 x_1, x_2 和 x_3 可能会比去计算 P^3 和 P^3x_0 更容易些. 但如果 P 比较小，比如说 30×30，在计算机上两种方法的计算时间都不长，我们应该首先计算 P^3x_0，这是因为它需要的人工输入较少.

基础练习

1. 假设某地区居民的迁移服从例题 1 迁移矩阵中的概率，“随机” 选择一个居民，则某一年的状态向量可以解释为那时他是城市居民或是郊区居民的概率.

295

a. 假设所选的居民现在是一个城市居民，则 $x_0=\begin{bmatrix}1\\0\end{bmatrix}$. 下一年他生活在郊区的可能性是多大?

b. 两年后他生活在郊区的可能性是多大?

2. 设 $P=\begin{bmatrix}0.6 & 0.2\\0.4 & 0.8\end{bmatrix}, q=\begin{bmatrix}0.3\\0.7\end{bmatrix}$, q 是否是 P 的稳态向量?

3. 例题 1 中，多年以后郊区的人口百分比会是多少?

习题　4.9

1. 某偏僻村庄可以接收到的无线电广播来自两个电台，一个新闻电台和一个音乐电台. 两个台的广播每隔半小时都会中断休息. 每当出现中断时，新闻台的听众有 70% 会继续收听新闻台，30% 的会换到音乐台；音乐台的听众有 60% 会换到新闻台，40% 会继续收听音乐台. 假设上午 8:15 所有人都在收听新闻台.
   a. 给出描述无线电听众在中断休息时换台的随机矩阵，并标注行和列.
   b. 给出初始的状态向量.
   c. 上午 9:25 音乐台的听众所占百分比是多少(假设电台在上午 8:30 和 9:00 中断休息)?
2. 一只实验室的动物每天可以吃三种食物中的任何一种. 实验室记录表明如果在一次实验中该动物选择了一种食物，则在下一次实验中，它有 50% 的概率选择同一种食物，选择其他两种食物的概率均为 25% .
   a. 这种情况的随机矩阵是怎样的?

b. 如果该动物在初始实验中选择了食物#1，则在之后的第 2 次实验中选择食物#2 的概率是多少?

3. 在任何给定的一天中，一个学生或是健康的或是生病的. 在今天所有的健康学生中，95%在第 2 天仍是健康的；在今天所有的生病的学生中，55% 在第 2 天仍是生病的.

a. 这种情况的随机矩阵是怎样的?

b. 假设在星期一有 20% 的学生生病，则在星期二生病的学生的百分率是多少? 星期三呢?

c. 如果一个学生是健康的，则从现在起两天后他或她仍是健康的概率是多少?

4. 在任何给定的一天中，哥伦布市的天气或者是好，或者是一般，或者是坏. 如果今天是好天气，则明天 60% 仍然是好天气，30% 变得一般，10% 变坏. 如果今天一般，则变好的概率是 0.40，仍然一般的概率是 0.30. 最后，如果今天是坏天气，则明天变好的概率是 0.40，变得一般的概率是 0.50.

a. 这种情况的随机矩阵是怎样的?

b. 假设今天是好天气或是一般的可能性均为 50%，则明天是坏天气的可能性是多大?

c. 假设预报星期一天气一般的概率是 40%，较坏的概率是 60%，则在星期三为好天气的可能性有多大?

求习题 5 ~ 8 中的稳态向量.

5. $\begin{bmatrix} 0.1 & 0.6 \\ 0.9 & 0.4 \end{bmatrix}$　6. $\begin{bmatrix} 0.8 & 0.5 \\ 0.2 & 0.5 \end{bmatrix}$　7. $\begin{bmatrix} 0.7 & 0.1 & 0.1 \\ 0.2 & 0.8 & 0.2 \\ 0.1 & 0.1 & 0.7 \end{bmatrix}$　8. $\begin{bmatrix} 0.7 & 0.2 & 0.2 \\ 0 & 0.2 & 0.4 \\ 0.3 & 0.6 & 0.4 \end{bmatrix}$

296

9. 判断矩阵 $\boldsymbol{P} = \begin{bmatrix} 0.2 & 1 \\ 0.8 & 0 \end{bmatrix}$是否是正则的随机矩阵.

10. 判断矩阵 $\boldsymbol{P} = \begin{bmatrix} 1 & 0.2 \\ 0 & 0.8 \end{bmatrix}$是否是正则的随机矩阵.

11. a. 求习题 1 中的马尔可夫链的稳态向量.

b. 在这一天的晚些时候，听新闻的听众的比例是多少?

12. 参考习题 2，动物经过多次选择后会更喜欢哪种食物?

13. a. 求习题 3 中的马尔可夫链的稳态向量.

b. 许多天以后某个学生生病的概率是多少? 今天是否生病对其有影响吗?

14. 参考习题 4，长时间看来，哥伦布市在给定的某一天为好天的可能性是多少?

15. [**M**]下面的移民矩阵由加州金融财政部人口统计研究联合会提供，它描述了 1989 年间美国人口的迁移情况. 1989 年，加利福尼亚的居民约占总人口的 11.7%. 如果所列的迁移概率常年保持不变，加利福尼亚的居民在总人口中的百分率最终将是多少?

从:		
加州	美国其他地方	到:
0.9821	0.0029	加州
0.0179	0.9971	美国其他地方

16. [**M**]在底特律，Hertz 汽车租赁公司有 2000 辆汽车. 下表中的比值给出了汽车出租和归还的情况. 在特定的某一天，市区内会有多少汽车被租，又有多少汽车准备出租?

车辆出租地:			
城市机场	市中心	大都会机场	归还到:
0.90	0.01	0.09	城市机场
0.01	0.90	0.01	市中心
0.09	0.09	0.90	大都会机场

17. 设 P 是一个 $n\times n$ 随机矩阵，下面的论述表明方程 $Px=x$ 有非平凡解(事实上,存在一个只含非负分量的稳态解,其证明可在本书后继教材找到). 为下列每个断言说明理由(需要时可引用定理).
    a. 如果 $P-I$ 的其他行都加到最末行，将得一个零行.
    b. $P-I$ 的各行是线性无关的.
    c. $P-I$ 的行空间的维数小于 n.
    d. $P-I$ 有非平凡零空间.
18. 证明：每个 2×2 随机矩阵至少有一个稳态向量. 任何一个这样的矩阵都可以写成 $P=\begin{bmatrix}1-\alpha & \beta\\ \alpha & 1-\beta\end{bmatrix}$，其中 α，β 是0，1之间的常数.（如果 $\alpha=\beta=0$,则有两个线性无关的稳态向量,否则只有一个）
19. 设 S 是每列均为1的 $1\times n$ 行矩阵,
    $S=[1\quad 1\quad \cdots\quad 1]$
    a. 请解释为什么 $\mathbf{R}^n$中的向量 x 是概率向量当且仅当它的分量是非负的并且 $Sx=1$.（1×1 矩阵,比如 Sx,在书写时通常不带方括号.）
    b. 设 P 是个 $n\times n$ 随机矩阵，请解释为什么 $SP=S$.
    c. 设 P 是个 $n\times n$ 随机矩阵，x 是概率向量，证明：Px 也是一个概率向量.
20. 利用习题19证明：如果 P 是个 $n\times n$ 随机矩阵，则 P^2 也是.
21. [M]分析正则随机矩阵的幂.
    a. 对于 $k=2$，3，4，5，计算 P^k，其中

$$P=\begin{bmatrix}0.3355 & 0.3682 & 0.3067 & 0.0389\\ 0.2663 & 0.2723 & 0.3277 & 0.5451\\ 0.1935 & 0.1502 & 0.1589 & 0.2395\\ 0.2047 & 0.2093 & 0.2067 & 0.1765\end{bmatrix}$$

    写出带有4位小数的计算结果，随着 k 的增加 P^k的列向量将如何变化？求出 P 的稳态向量.
    b. 对于 $k=10,20,\cdots,80$,计算 Q^k，其中

$$Q=\begin{bmatrix}0.97 & 0.05 & 0.10\\ 0 & 0.90 & 0.05\\ 0.03 & 0.05 & 0.85\end{bmatrix}$$

    （要使 Q^k精确到小数点后4位,可能需要 $k=116$ 或者更大)，计算 Q 的稳态向量. 猜测对任意正则随机矩阵，结果可能是怎样的.

    297

    c. 利用定理18来解释你在(a)和(b)中的发现.
22. [M]对比求正则随机矩阵 P 的稳态向量 q 的两种方法：(1)像例题5中那样求出 q，或者(2)对于某个足够大的 k 值求 P^k，然后将 P^k中的一个列向量作为 q 的近似值. [学习指南中介绍的 nulbasis 程序基本能够自动实现方法(1).]

    利用你的矩阵软件所允许的最高阶随机矩阵，并且取 $k=100$ 或者更大，来验证你的程序. 对每一种方法，说明键盘输入和程序运行的总时间(MATLAB 的某些版本已经具有命令 flops 和 tic…toc,它们分别用来记浮点运算的次数和 MATLAB 用去的总时间). 比较每种办法的优缺点，并声明你更喜欢哪一种.

基础练习答案

1. a. 在一年内，5%的城市居民将迁移到郊区，所以选择这个人的可能性是5%. 我们没有关

于这个人的其他信息，因此说他迁往郊区的可能性为 5%．这一事实包含于状态向量 x_1 的第 2 个分量中，其中

$$x_1 = Mx_0 = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.95 \\ 0.05 \end{bmatrix}$$

b. 两年后他生活在郊区的可能性为 9.6%，这是由于

$$x_2 = Mx_1 = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix}\begin{bmatrix} 0.95 \\ 0.05 \end{bmatrix} = \begin{bmatrix} 0.904 \\ 0.096 \end{bmatrix}$$

2. 稳态向量应满足 $Px = x$，因为

$$Pq = \begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix}\begin{bmatrix} 0.3 \\ 0.7 \end{bmatrix} = \begin{bmatrix} 0.32 \\ 0.68 \end{bmatrix} \neq q$$

我们断定 q 不是 P 的稳态向量．

3. 例题 1 中的 M 是正则随机矩阵，因为它的分量全部是严格正的．所以我们可以运用定理 18．从例题 4 中我们已经得到稳态向量，因此人口分布向量 x_k 收敛于

$$q = \begin{bmatrix} 0.375 \\ 0.625 \end{bmatrix}$$

最终将有 62.5% 的人口生活在郊区．

第 4 章补充题

1. 判断命题的真假，并说明理由（如果命题为真，引用恰当的事实或定理．如果为假，解释原因或举一反例说明命题不能在每一种情况下都正确）．(a)～(f) 中，$v_1, \cdots, v_p$ 是一个非零有限维向量空间中的向量，并且 $S = \{v_1, \cdots, v_p\}$．

   a. $v_1, \cdots, v_p$ 的全体线性组合所构成的集合是一个向量空间．

   b. 如果 $\{v_1, \cdots, v_{p-1}\}$ 张成 V，则 S 张成 V．

   c. 如果 $\{v_1, \cdots, v_{p-1}\}$ 是线性无关的，则 S 也是．

   d. 如果 S 是线性无关的，则 S 是 V 的一组基．

   e. 如果 Span $S = V$，则 S 的某个子空间是 V 的一组基．

   f. 如果 dim $V = p$，并且 Span $S = V$，则 S 不可能是线性相关的．

   g. $\mathbf{R}^3$ 中的平面是 2 维子空间．

   h. 矩阵中的非主元列总是线性相关的．

   i. 对矩阵 A 进行行变换能够改变 A 的行之间的线性相关关系．

   j. 矩阵的行变换会改变零空间．

   k. 矩阵的秩等于非零行的个数．

   l. 如果 $m \times n$ 矩阵 A 与阶梯矩阵 U 行等价，并且 U 有 k 个非零行，则 $Ax = 0$ 的解空间的维数是 $m - k$．

   m. 如果 B 是矩阵 A 通过几次初等行变换得到的，则 rank B = rank A．

   n. 矩阵 A 的非零行构成 Row A 的一组基．

   o. 如果矩阵 A，B 有相同的化简后的阶梯矩阵，则 Row A = Row$B$．

   p. 如果 H 是 $\mathbf{R}^3$ 的一个子空间，则存在一个 3×3 矩阵 A 使得 H = Col A．

   q. 如果 A 是 $m \times n$ 矩阵，并且 rank $A = m$，则线性变换 $x \to Ax$ 是一对一的．

   r. 如果 A 是 $m \times n$ 的，并且线性变换 $x \to Ax$ 是到上的，则 rank $A = m$．

298

s. 坐标变换矩阵总是可逆的.

t. 如果 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$和 $\mathcal{C}=\{\boldsymbol{c}_1,\cdots,\boldsymbol{c}_n\}$是向量空间的基，则坐标变换矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ 的第 j 列是坐标向量$[\boldsymbol{c}_j]_{\mathcal{B}}$.

2. 求所有下面形式的向量构成的集合的基

$$\begin{bmatrix} a-2b+5c \\ 2a+5b-8c \\ -a-4b+7c \\ 3a+b+c \end{bmatrix}.\text{（计算时要仔细）}$$

3. 设 $\boldsymbol{u}_1=\begin{bmatrix} -2 \\ 4 \\ -6 \end{bmatrix},\boldsymbol{u}_2=\begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix},\boldsymbol{b}=\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$，$W=\text{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$. 求 W 的一个隐式表示，即用一个或多个齐次方程来刻画 W 中的点. [提示：$\boldsymbol{b}$ 何时在 W 中?]

4. 请解释下列论证为什么是错误的：设$\boldsymbol{f}(t)=3+t,\boldsymbol{g}(t)=3t+t^2$，并且注意到 $\boldsymbol{g}(t)=t\boldsymbol{f}(t)$. 则有$\{\boldsymbol{f},\boldsymbol{g}\}$是线性相关的，这是由于 $\boldsymbol{g}$ 是$\boldsymbol{f}$的倍数.

5. 考虑多项式$\boldsymbol{p}_1(t)=1+t,\boldsymbol{p}_2(t)=1-t,\boldsymbol{p}_3(t)=4,\boldsymbol{p}_4(t)=t+t^2,\boldsymbol{p}_5(t)=1+2t+t^2$，并且设 H 是由 $S=\{\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3,\boldsymbol{p}_4,\boldsymbol{p}_5\}$张成的 $\mathbf{P}_5$的子空间. 利用张成集定理(4.3节)证明中描述的方法来得到 H 的一组基(解释如何选择 S 中恰当的元素).

6. 假设$\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3,\boldsymbol{p}_4$是张成 $\mathbf{P}_5$的2维子空间的具体的多项式. 描述如何通过验证这4个多项式和进行尽可能少的运算来求 H 的一组基.

7. 对于一个含18个线性方程和20个变量的齐次方程组，为了保证它的每个相关的非齐次方程都有解，齐次方程组的解集应满足什么性质? 试加以讨论.

8. 设 H 是 n 维向量空间 V 的 n 维子空间，请解释为什么 $H=V$.

9. 设 T：$\mathbf{R}^n\to\mathbf{R}^n$是个线性变换.

   a. 如果 T 是个一对一映射，则它的值与它的维数是多少? 请解释.

   b. 如果 T 将 $\mathbf{R}^n$映射到 $\mathbf{R}^m$上，则 T 的核(见4.2节)的维数是多少? 请解释.

10. 设 S 是向量空间 V 的一个极大线性无关子集，即 S 具有下列性质：把一个不属于 S 的向量加入到 S 中得到的新集合不再线性无关. 证明：S 一定是 V 的一组基. [提示：如果 S 是线性无关但不是 V 的一组基会怎样?]

11. 设 S 是向量空间 V 的一个有限极小张成集，即 S 具有性质：从 S 去掉一个向量得到的新集合不再能张成 V. 证明：S 一定是 V 的一组基.

习题12～17推广了在应用中有时会用到的秩的性质，假设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵.

12. 通过(a)和(b)证明：$\boldsymbol{AB}$ 的秩不能超过 $\boldsymbol{A}$ 或 $\boldsymbol{B}$ 的秩(一般地,矩阵的积的秩不能超过任何一个因子的秩).

    a. 证明：如果 $\boldsymbol{B}$ 是 $n\times p$ 矩阵，则 $\text{rank}\,\boldsymbol{AB}\leqslant\text{rank}\,\boldsymbol{A}$. [提示：解释为什么 $\boldsymbol{AB}$ 的列空间中的每个向量都在 $\boldsymbol{A}$ 的列空间中.]

    b. 证明：如果 $\boldsymbol{B}$ 是 $n\times p$ 矩阵，则 $\text{rank}\,\boldsymbol{AB}\leqslant\text{rank}\,\boldsymbol{B}$. [提示：利用(a)来研究 $\text{rank}(\boldsymbol{AB})^{\text{T}}$.]

13. 证明：如果 $\boldsymbol{P}$ 是可逆的 $m\times m$ 矩阵，则 $\text{rank}\boldsymbol{PA}=\text{rank}\,\boldsymbol{A}$. [提示：把习题12应用到 $\boldsymbol{PA}$ 和 $\boldsymbol{P}^{-1}(\boldsymbol{PA})$上.]

299

14. 证明：如果 $\boldsymbol{Q}$ 是可逆的，则 $\text{rank}\,\boldsymbol{AQ}=\text{rank}\,\boldsymbol{A}$. [提示：利用习题13来研究 $\text{rank}(\boldsymbol{AQ})^{\text{T}}$.]

15. 设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，$\boldsymbol{B}$ 是 $n\times p$ 矩阵，并且满足 $\boldsymbol{AB}=\boldsymbol{0}$，证明：$\text{rank}\,\boldsymbol{A}+\text{rank}\,\boldsymbol{B}\leqslant n$. [提示：$\text{Nul}\,\boldsymbol{A}$，$\text{Col}\,\boldsymbol{A}$，$\text{Nul}\,\boldsymbol{B}$，$\text{Col}\,\boldsymbol{B}$ 中的一个包含在其他3个子空间的一个中.]

16. 如果 $\boldsymbol{A}$ 是个秩为 r 的 $m\times n$ 矩阵，则 $\boldsymbol{A}$ 的秩分解为等式 $\boldsymbol{A}=\boldsymbol{CR}$，其中 $\boldsymbol{C}$ 是秩为 r 的 $m\times r$

矩阵，$\boldsymbol{R}$ 是秩为 r 的 $r\times n$ 矩阵，这样的分解总是存在(4.6 节习题 38)．给定任何两个 $m\times n$矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$，利用 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的秩分解来证明：

$$\text{rank }(\boldsymbol{A}+\boldsymbol{B})\leqslant \text{rank }\boldsymbol{A}+\text{rank }\boldsymbol{B}$$

[提示：将 $\boldsymbol{A}+\boldsymbol{B}$ 写成两个矩阵的积．]

17. 矩阵 $\boldsymbol{A}$ 的一个**子矩阵**(submarix)是从$\boldsymbol{A}$ 中删除某些行或者列后得到的结果．可以证明，$\boldsymbol{A}$ 的秩为 r 当且仅当 $\boldsymbol{A}$ 含有一个可逆的 $r\times r$ 子矩阵并且任何更大的子矩阵不可逆．通过解释下面的 a 和 b 来证明该命题．
   a. 为什么秩为 r 的 $m\times n$ 矩阵 $\boldsymbol{A}$ 含有一个秩为 r 的 $m\times r$ 子矩阵 $\boldsymbol{A}_1$？
   b. 为什么 $\boldsymbol{A}_1$ 含有一个可逆的 $r\times r$ 的子矩阵 $\boldsymbol{A}_2$？

秩的概念在工程控制系统有着非常重要的作用，比如在本章实例介绍中提到的航天飞机系统中，控制系统的一个状态空间包含下面形式的差分方程

$$\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k+\boldsymbol{B}\boldsymbol{u}_k \quad 对于\ k=0,1,\cdots \tag{1}$$

其中 $\boldsymbol{A}$ 是 $n\times n$ 矩阵，$\boldsymbol{B}$ 是 $n\times m$ 矩阵，$\{\boldsymbol{x}_k\}$是 $\mathbf{R}^n$ 中的一列“状态向量”，它描述了系统在离散时间的状态，$\{\boldsymbol{u}_k\}$是控制或输出序列．称$(\boldsymbol{A},\boldsymbol{B})$是**可控的**(controllable)，如果

$$\text{rank}[\boldsymbol{B}\quad \boldsymbol{AB}\quad \boldsymbol{A}^2\boldsymbol{B}\quad \cdots\quad \boldsymbol{A}^{n-1}\boldsymbol{B}]=n \tag{2}$$

称(2)中的矩阵为系统的**可控矩阵**(controllability matrix)．如果$(\boldsymbol{A},\boldsymbol{B})$是可控的，则说该系统可控制，即，通过在 $\mathbf{R}^m$中选择恰当的控制序列，就可以从状态 $\mathbf{0}$ 开始在至多 n 步内到达指定的某个状态 $\boldsymbol{v}$(在 $\mathbf{R}^n$ 中)．习题 18 表明，当 $n=4$ 和 $m=2$ 时，该事实是成立的．关于可控性的进一步讨论，见课本的网站(第 4 章的案例研究．)

18. 假设 $\boldsymbol{A}$ 是 4×4 矩阵，$\boldsymbol{B}$ 是 4×2 矩阵，并且设 $\boldsymbol{u}_0,\cdots,\boldsymbol{u}_3$代表 $\mathbf{R}^2$ 中的一列输入向量．
   a. 令 $\boldsymbol{x}_0=\mathbf{0}$，根据(1)计算 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_4$，并写出 $\boldsymbol{x}_4$的含有在(2)中出现的可控矩阵 $\boldsymbol{M}$ 的公式．
   b. 假设$(\boldsymbol{A},\boldsymbol{B})$是可控的，$\boldsymbol{v}$ 是 $\mathbf{R}^4$ 中的任意向量，解释为什么存在 $\mathbf{R}^2$ 中的控制序列 $\boldsymbol{u}_0,\cdots,\boldsymbol{u}_3$使得 $\boldsymbol{x}_4=\boldsymbol{v}$．

判断习题 19～22 中的矩阵对是否可控．

19. $\boldsymbol{A}=\begin{bmatrix}0.9 & 1 & 0\\ 0 & -0.9 & 0\\ 0 & 0 & 0.5\end{bmatrix}$，$\boldsymbol{B}=\begin{bmatrix}0\\1\\1\end{bmatrix}$

20. $\boldsymbol{A}=\begin{bmatrix}0.8 & -0.3 & 0\\ 0.2 & 0.5 & 1\\ 0 & 0 & -0.5\end{bmatrix}$，$\boldsymbol{B}=\begin{bmatrix}1\\1\\0\end{bmatrix}$

21. [M]$\boldsymbol{A}=\begin{bmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ -2 & -4.2 & -4.8 & -3.6\end{bmatrix}$，$\boldsymbol{B}=\begin{bmatrix}1\\0\\0\\-1\end{bmatrix}$

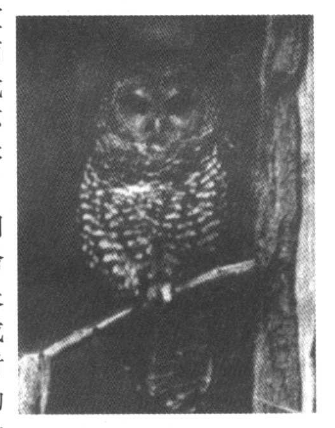
22. [M]$\boldsymbol{A}=\begin{bmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ -1 & -13 & -12.2 & -1.5\end{bmatrix}$，$\boldsymbol{B}=\begin{bmatrix}1\\0\\0\\-1\end{bmatrix}$

300

第5章　本征值与本征向量

实例介绍：动态系统和斑点猫头鹰

1990年，对于是否使用太平洋西北岸茂密森林中的木材，人们展开了广泛的争论，争论的焦点就是北方斑点猫头鹰.



环保主义者使联邦政府确信，如果继续砍伐那片适宜猫头鹰生活的原始森林(树龄超过200年)，猫头鹰将面临灭绝的威胁．木材产业者预计，如果政府采取了限制伐木的措施，将会有30 000到100 000人失业，他们认为不应该把猫头鹰划入“濒临灭绝的物种”，并且引用了许多已发表的科学报道支持这一论断.[1]

数学生态学者们身处两派的争论中，这更激发了他们了解斑点猫头鹰种群动态的热情．一只斑点猫头鹰的生命周期可以自然地分为三个阶段：幼鸟期(1岁以下)、成长期(1到2岁)和成熟期(2岁以上)．猫头鹰在成长期和成熟期交配，从成熟期开始繁殖，最长可活至20岁．每对猫头鹰大约需要1000公顷(4平方英里)作为它们自己的领地．生命周期中很重要的一个时期就是幼鸟离开巢穴的这段过程．为了存活下去并继续生长，它必须成功找到一个属于自己的新领地(通常还有一个配偶).

研究猫头鹰种群动态的第一步就是以年为时间间隔来建模，记时间为 $k=0,1,2,\cdots$，通常假设在各个生命阶段雌雄猫头鹰的数量之比为1∶1，于是可以只计算雌猫头鹰的数量．我们用向量 $\boldsymbol{x}_k=(j_k,s_k,a_k)$ 来表示在第 k 年雌猫头鹰的数量，其中 j_k，S_k，a_k 分别表示处于幼鸟期、成长期和成熟期的数量.

301

根据人口统计学研究中的实际野外数据，R. Lamberson和他的同事们建立了下列状态－矩阵模型：[2]

$$\begin{bmatrix} j_{k+1} \\ s_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0.33 \\ 0.18 & 0 & 0 \\ 0 & 0.71 & 0.94 \end{bmatrix} \begin{bmatrix} j_k \\ s_k \\ a_k \end{bmatrix}$$

从中可以看出，基于每一对猫头鹰的平均生育率，第 $k+1$ 年幼鸟期雌猫头鹰的

---

1. “The Great Spotted Owl War,” *Reader's Digest*, November 1992, pp. 91～95.
2. R. H. Lamberson, R. McKelvey, B. R. Noon, and C. Voss, “A Dynamic Analysis of the Viability of the Northern Spotted Owl in a Fragmented Forest Environment,” *Conservation Biology* 6(1992), 505～512. 也可参考Lamberson教授1993年的一封私人信件.

数量是第 k 年成熟期雌猫头鹰数量的 0.33 倍. 此外, 有 18% 的幼鸟能够存活下来进入成长期; 有 71% 的处于成长期的猫头鹰和 94% 的处于成熟期的猫头鹰能够存活下来, 且被计入第 $k+1$ 年成熟期的猫头鹰之列.

上述状态-矩阵模型是一个形如 $\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k$ 的差分方程. 因为它描述了系统随时间的变化, 因此这样的方程通常称作**动态系统**(dynamical system)或**离散线性动态系统**(discrete linear dynamical system).

Lamberson 状态矩阵的各个元素中, 幼鸟 18% 的存活率主要受现存原始森林数量的影响. 事实上, 通常有 60% 的幼鸟可以存活并离开旧巢, 但在 Lamberson 和他的同事们所研究的加利福尼亚柳河湾地区, 离开旧巢的幼鸟中只有 30% 可以成功找到自己的新领地, 其余的都在寻找的过程中死亡了.

猫头鹰不能成功地找到新领地的一个主要原因是过多的皆伐区使这片原始森林变得支离破碎, 并且这种破坏仍在继续. 当猫头鹰离开森林树冠的保护穿越皆伐区时, 被捕食者攻击的危险急剧增加. 5.6 节中将说明, 上述模型预示了猫头鹰最后必将灭绝. 不过, 如果有 50% 的幼鸟能够存活下来离开旧巢并找到新领地, 那么猫头鹰种群就会兴旺.

本章的目标是将线性变换 $\boldsymbol{x}\mapsto\boldsymbol{A}\boldsymbol{x}$ 的作用分解成易于理解的成分. 除 5.4 节外, 本章所讨论的矩阵均为方阵. 这里介绍的主要应用是离散动态系统, 包括前面讨论的斑点猫头鹰. 不过本征值和本征向量这两个基本概念在纯数学和应用数学中都有广泛的应用, 并且它们出现的背景也比我们这里考虑的要一般得多. 本征值还可用于研究微分方程和连续动态系统, 它们为工程设计提供了重要的信息, 甚至自然而然地出现在物理、化学等领域.

5.1 本征向量与本征值

尽管变换 $\boldsymbol{x}\mapsto\boldsymbol{A}\boldsymbol{x}$ 可能会把向量往各种方向上移动, 但往往存在一些特殊的向量, $\boldsymbol{A}$ 在其上的作用十分简单.

【例题 1】 设 $\boldsymbol{A}=\begin{bmatrix}3 & -2\\1 & 0\end{bmatrix}, \boldsymbol{u}=\begin{bmatrix}-1\\1\end{bmatrix}, \boldsymbol{v}=\begin{bmatrix}2\\1\end{bmatrix}$. $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 在 $\boldsymbol{A}$ 作用下的像如图5-1所示. 事实上, $\boldsymbol{A}\boldsymbol{v}$ 就是 $2\boldsymbol{v}$. 所以, $\boldsymbol{A}$ 只是拉伸了 $\boldsymbol{v}$. ■

302

作为另一个例子, 读者可以回忆 4.9 节: 如果 $\boldsymbol{A}$ 是一个随机矩阵, 则 $\boldsymbol{A}$ 的稳态向量 $\boldsymbol{q}$ 满足方程 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}$, 也就是 $\boldsymbol{A}\boldsymbol{q}=1\cdot\boldsymbol{q}$.

在本节中, 我们研究形如

$$\boldsymbol{A}\boldsymbol{x}=2\boldsymbol{x} \text{ 或 } \boldsymbol{A}\boldsymbol{x}=-4\boldsymbol{x}$$

的方程, 并且求那些被 $\boldsymbol{A}$ 作用相当于被数乘作用的向量。

【定义】 $\boldsymbol{A}$ 为 $n\times n$ 矩阵, 对非零向量 $\boldsymbol{x}$, 若存在某个数量 λ, 使得 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$, 则称 $\boldsymbol{x}$ 为 $\boldsymbol{A}$ 的**本征向量**(eigenvector). 如果方程 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$ 有非平凡解 $\boldsymbol{x}$, 则称数量 λ 为 $\boldsymbol{A}$ 的**本征值**(eigenvalue); 这样的 $\boldsymbol{x}$ 称为对应于本征值 λ 的本征向量.[1]

1. 注意, 根据定义本征向量必须非零, 但本征值可能为零. 数量 0 是本征值的情况将在例题 5 后面讨论.

容易判定一个给定的向量是否为一个矩阵的本征向量以及一个指定的数量是否为本征值.

【例题2】　设 $A=\begin{bmatrix}1&6\\5&2\end{bmatrix}, u=\begin{bmatrix}6\\-5\end{bmatrix}, v=\begin{bmatrix}3\\-2\end{bmatrix}$. 判断 u 和 v 是否为 A 的本征值.

解: $Au=\begin{bmatrix}1&6\\5&2\end{bmatrix}\begin{bmatrix}6\\-5\end{bmatrix}=\begin{bmatrix}-24\\20\end{bmatrix}=-4\begin{bmatrix}6\\-5\end{bmatrix}=-4u$

$$Av=\begin{bmatrix}1&6\\5&2\end{bmatrix}\begin{bmatrix}3\\-2\end{bmatrix}=\begin{bmatrix}-9\\11\end{bmatrix}\neq\lambda\begin{bmatrix}3\\-2\end{bmatrix}$$

所以 u 是对应于某个本征值(-4)的本征向量，但 v 不是 A 的本征向量，因为 Av 不是 v 的倍数. ■

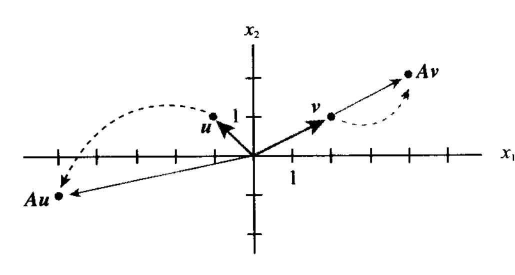


图5-1　乘以 A 的效果

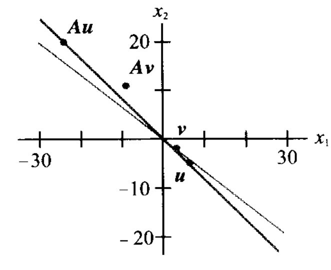


图5-2　$Au=-4u$，但 $Av\neq\lambda v$

【例题3】　证明7是例题2中矩阵 A 的本征值，并求本征值7对应的本征向量.

解: 7为 A 的本征值当且仅当方程

303

$$Ax=7x \tag{1}$$

有非平凡解. 而(1)等价于 $Ax-7x=0$,或

$$(A-7I)x=0 \tag{2}$$

为求解该齐次方程,构造如下矩阵

$$A-7I=\begin{bmatrix}1&6\\5&2\end{bmatrix}-\begin{bmatrix}7&0\\0&7\end{bmatrix}=\begin{bmatrix}-6&6\\5&-5\end{bmatrix}$$

显然，$A-7I$ 的列向量线性相关，所以(2)有非平凡解. 因此7是 A 的本征值. 为求出7对应的本征向量，做行变换:

$$\begin{bmatrix}-6&6&0\\5&-5&0\end{bmatrix}\sim\begin{bmatrix}1&-1&0\\0&0&0\end{bmatrix}$$

其通解为 $x_2\begin{bmatrix}1\\1\end{bmatrix}$. 对任意 $x_2\neq0, x_2\begin{bmatrix}1\\1\end{bmatrix}$都为 A 的对应于本征值7的本征向量. ■

警告:尽管在例题3中使用了行化简来求解本征向量,但不能用这种方法来求解特征值. 这是因为矩阵 A 的阶梯形式通常不能反映 A 的本征值.

显然,以任意 λ 代替 $\lambda=7$,方程(1)和(2)都等价. 于是，λ 是 A 的本征值当且仅当方程

$$(A-\lambda I)x=0 \tag{3}$$

有非平凡解.(3)所有解的集合恰好是矩阵 $A-\lambda I$ 的零空间. 所以这个集合是 $\mathbf{R}^n$ 的子空间,我们将它称为 A 的对应于 λ 的**本征空间**(eigenspace). 这个本征空间包含零向量和全体对应于本征值 λ 的本征向量.

从例题3可看出,对于例题2中的矩阵 A,7对应的本征空间由(1,1)的全体倍数构成,即过(1,1)和原点的直线. 对于例题2,可以验证本征值 $\lambda=-4$ 对应的本征空间是通过(6,-5)的直线. 这些本征空间与本征向量(1,1),(3/2,-5/4)以及变换 $x\mapsto Ax$ 在每个本征空间上的几何作用一起表示在图5-3中. ■

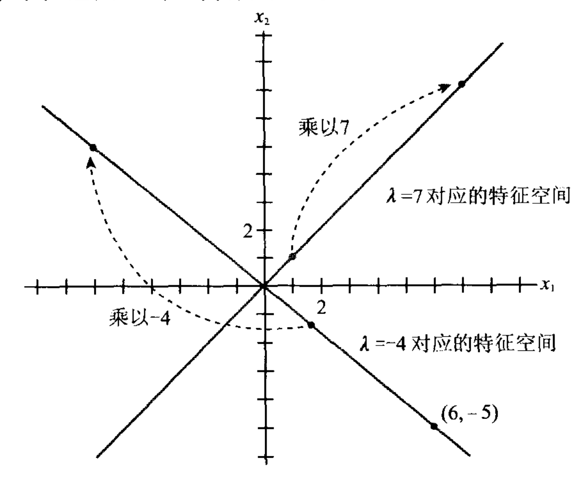


图5-3 $\lambda=-4$ 和 $\lambda=7$ 的本征空间

【例题4】 设 $A=\begin{bmatrix}4&-1&6\\2&1&6\\2&-1&8\end{bmatrix}$. A 的一个本征值为2. 求对应于2的本征空间的一组基.

解:构造

$$A-2I=\begin{bmatrix}4&-1&6\\2&1&6\\2&-1&8\end{bmatrix}-\begin{bmatrix}2&0&0\\0&2&0\\0&0&2\end{bmatrix}=\begin{bmatrix}2&-1&6\\2&-1&6\\2&-1&6\end{bmatrix}$$

对 $(A-2I)x=\mathbf{0}$ 的增广矩阵进行行化简:

$$\begin{bmatrix}2&-1&6&0\\2&-1&6&0\\2&-1&6&0\end{bmatrix}\sim\begin{bmatrix}2&-1&6&0\\0&0&0&0\\0&0&0&0\end{bmatrix}$$

304

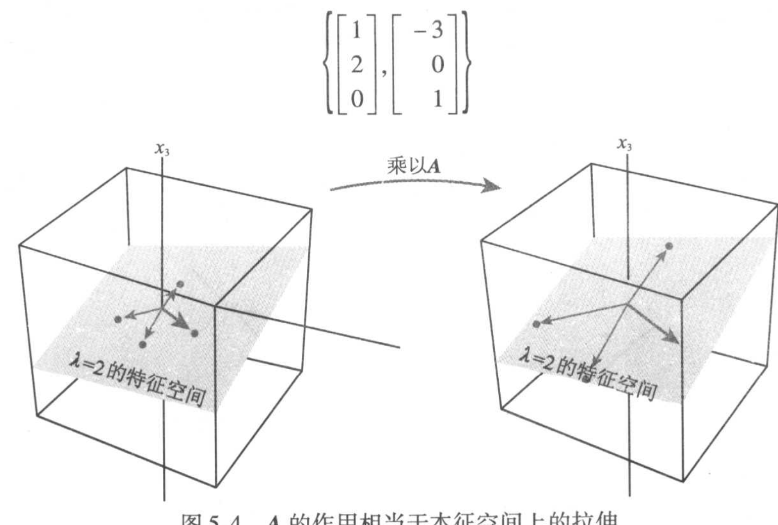
因为方程 $(A-2I)x=\mathbf{0}$ 有自由变量,所以我们可以断定2确实为 A 的本征值. 通解为

$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=x_2\begin{bmatrix}1/2\\1\\0\end{bmatrix}+x_3\begin{bmatrix}-3\\0\\1\end{bmatrix},\text{ 其中 }x_2,\ x_3\text{ 为自由变量}$$

如图5-4所示,这个本征空间为 $\mathbf{R}^3$ 的一个二维子空间. 它的一组基为

$$\left\{\begin{bmatrix}1\\2\\0\end{bmatrix},\begin{bmatrix}-3\\0\\1\end{bmatrix}\right\}$$

■



305

图5-4　A 的作用相当于本征空间上的拉伸

注记

在已知本征值这种简单情况下，例题4给出了一种手算本征向量的好方法. 使用矩阵软件和行化简来求(某一指定本征值对应的)本征空间通常也奏效，但并不完全可靠. 舍入误差有时会导致得到一个有错误主元个数的阶梯形矩阵. 对于那些不太大的矩阵，最好的计算机程序可以同时计算出本征值和本征向量的近似值，并且可以达到任意指定的精度. 随着计算机计算性能和软件的不断改进，我们能够分析处理越来越大规模的矩阵.

在为数不多的几种特殊情况下，本征值可以被精确地求出，下面的定理就属于其中一种情况. 关于本征值的计算还将在5.2节中继续讨论.

【定理1】　三角矩阵的本征值为其主对角线上的元素.

证明： 简单起见，考虑3×3的情形. 若 A 为上三角矩阵，计算 $A-\lambda I$ 如下：

$$A-\lambda I=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\0&a_{22}&a_{23}\\0&0&a_{33}\end{bmatrix}-\begin{bmatrix}\lambda&0&0\\0&\lambda&0\\0&0&\lambda\end{bmatrix}=\begin{bmatrix}a_{11}-\lambda&a_{12}&a_{13}\\0&a_{22}-\lambda&a_{23}\\0&0&a_{33}-\lambda\end{bmatrix}$$

数量 λ 为 A 的本征值当且仅当方程 $(A-\lambda I)x=0$ 有非平凡解，即当且仅当方程有自由变量. 由 $A-\lambda I$ 中的零元素，容易看出 $(A-\lambda I)x=0$ 有自由变量当且仅当 $A-\lambda I$ 的对角线上至少有一元素为零. 而这当且仅当 λ 等于 a_{11}，a_{22}，a_{33} 其中之一. 对于 A 是下三角矩阵的情况，见习题28. ■

【例题5】　设 $A=\begin{bmatrix}3&6&-8\\0&0&6\\0&0&2\end{bmatrix}$，$B=\begin{bmatrix}4&0&0\\-2&1&0\\5&3&4\end{bmatrix}$. A 的本征值为3，0和2. B 的本征值为4和1. ■

一个矩阵 A 有一个本征值为0，比如例题5，将意味着什么？0是 A 的本征值当

且仅当方程

$$Ax = 0x \tag{4}$$

有非平凡解. 而(4)等价于 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$，该方程有非平凡解当且仅当 $\boldsymbol{A}$ 不可逆. 所以 0 是 $\boldsymbol{A}$ 的本征值当且仅当 $\boldsymbol{A}$ 不可逆. 这个事实将被纳入 5.2 节的可逆矩阵定理.

306

下述重要定理以后将会用到. 它的证明是一个关于本征向量的典型计算.

【定理 2】 若 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r$ 为 $n\times n$ 矩阵 $\boldsymbol{A}$ 的对应于不同本征值 $\lambda_1,\cdots,\lambda_r$ 的本征向量，则集合$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r\}$线性无关.

证明： 假设$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r\}$线性相关. 因为 $\boldsymbol{v}_1$ 非零，由 1.7 节定理 7 知，有一个向量是其前面向量的线性组合. 记 p 为使 $\boldsymbol{v}_{p+1}$ 为其前面(线性无关)向量线性组合的最小的指标，则存在 $c_1,\cdots,c_p$ 使得

$$c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p=\boldsymbol{v}_{p+1} \tag{5}$$

因为对每个 k,都有 $\boldsymbol{A}\boldsymbol{v}_k=\lambda_k\boldsymbol{v}_k$，(5)两边同乘 A，得到

$$c_1\boldsymbol{A}\boldsymbol{v}_1+\cdots+c_p\boldsymbol{A}\boldsymbol{v}_p=\boldsymbol{A}\boldsymbol{v}_{p+1}$$

$$c_1\lambda_1\boldsymbol{v}_1+\cdots+c_p\lambda_p\boldsymbol{v}_p=\lambda_{p+1}\boldsymbol{v}_{p+1} \tag{6}$$

(5)两边同乘 λ_{p+1}后减去(6)，得到

$$c_1(\lambda_1-\lambda_{p+1})\boldsymbol{v}_1+\cdots+c_p(\lambda_p-\lambda_{p+1})\boldsymbol{v}_p=\boldsymbol{0} \tag{7}$$

因为$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$线性无关，所以(7)中系数都为 0. 又因为本征值互不相等，所以每个 $\lambda_i-\lambda_{p+1}$都不为零. 所以对 $i=1,\cdots,p$ 都有 $c_i=0$. 由(5)得到 $\boldsymbol{v}_{p+1}=\boldsymbol{0}$，矛盾. 因此$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r\}$线性无关. ■

本征向量和差分方程

我们现在来说明如何构造本章实例介绍所讨论的一阶差分方程的解，并以此结束本节.

$$\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k\ (k=0,1,2,\cdots) \tag{8}$$

如果 $\boldsymbol{A}$ 为 $n\times n$ 矩阵，则(8)是 $\mathbf{R}^n$ 中序列$\{\boldsymbol{x}_k\}$的一个*递归*描述. (8)的一个**解**(solution)是$\{\boldsymbol{x}_k\}$的一个显式表示，它对每个$\boldsymbol{x}_k$给出的公式既不直接依赖于 $\boldsymbol{A}$，也不依赖于序列当中$\boldsymbol{x}_k$前面除初始项 $\boldsymbol{x}_0$ 以外的那些项.

构造(8)的解最简单的方法是取一个本征向量 $\boldsymbol{x}_0$ 以及 $\boldsymbol{x}_0$ 所对应的本征值 λ，令

$$\boldsymbol{x}_k=\lambda^k\boldsymbol{x}_0\ (k=1,2,\cdots) \tag{9}$$

该序列即为(8)的解，这是因为

$$\boldsymbol{A}\boldsymbol{x}_k=\boldsymbol{A}(\lambda^k\boldsymbol{x}_0)=\lambda^k(\boldsymbol{A}\boldsymbol{x}_0)=\lambda^k(\lambda\boldsymbol{x}_0)=\lambda^{k+1}\boldsymbol{x}_0=\boldsymbol{x}_{k+1}$$

形如(9)的解的线性组合也是(8)的解！见习题 33.

307

基础练习

1. 5 是 $\boldsymbol{A}=\begin{bmatrix}6&-3&1\\3&0&5\\2&2&6\end{bmatrix}$的本征值吗？

2. 若 $\boldsymbol{x}$ 是 $\boldsymbol{A}$ 的对应于本征值 λ 的本征向量，$\boldsymbol{A}^3\boldsymbol{x}$ 是多少？

习题 5.1

1. $\lambda=2$ 是$\begin{bmatrix}3&2\\3&8\end{bmatrix}$的本征值吗？为什么？

2. $\lambda=-2$ 是$\begin{bmatrix}7 & 3\\3 & -1\end{bmatrix}$的本征值吗？为什么？

3. $\begin{bmatrix}1\\4\end{bmatrix}$ 是$\begin{bmatrix}-3 & 1\\-3 & 8\end{bmatrix}$的本征向量吗？如果是，求出对应的本征值.

4. $\begin{bmatrix}-1+\sqrt{2}\\1\end{bmatrix}$是$\begin{bmatrix}2 & 1\\1 & 4\end{bmatrix}$的本征向量吗？如果是，求出对应的本征值.

5. $\begin{bmatrix}4\\-3\\1\end{bmatrix}$是$\begin{bmatrix}3 & 7 & 9\\-4 & -5 & 1\\2 & 4 & 4\end{bmatrix}$的本征向量吗？如果是，求出对应的本征值.

6. $\begin{bmatrix}1\\-2\\1\end{bmatrix}$是$\begin{bmatrix}3 & 6 & 7\\3 & 3 & 7\\5 & 6 & 5\end{bmatrix}$的本征向量吗？如果是，求出对应的本征值.

7. $\lambda=4$ 是$\begin{bmatrix}3 & 0 & -1\\2 & 3 & 1\\-3 & 4 & 5\end{bmatrix}$的本征值吗？如果是，求出对应于 $\lambda=4$ 的本征向量.

8. $\lambda=3$ 是$\begin{bmatrix}1 & 2 & 2\\3 & -2 & 1\\0 & 1 & 1\end{bmatrix}$的本征值吗？如果是，求出对应于 $\lambda=3$ 的本征向量.

在习题9~16中，求每个给定本征值所对应的本征空间的一组基.

9. $\boldsymbol{A}=\begin{bmatrix}5 & 0\\2 & 1\end{bmatrix}$，$\lambda=1$，5

10. $\boldsymbol{A}=\begin{bmatrix}10 & -9\\4 & -2\end{bmatrix}$，$\lambda=4$

11. $\boldsymbol{A}=\begin{bmatrix}4 & -2\\-3 & 9\end{bmatrix}$，$\lambda=10$

12. $\boldsymbol{A}=\begin{bmatrix}7 & 4\\-3 & -1\end{bmatrix}$，$\lambda=1$，5

13. $\boldsymbol{A}=\begin{bmatrix}4 & 0 & 1\\-2 & 1 & 0\\-2 & 0 & 1\end{bmatrix}$，$\lambda=1$，2，3

14. $\boldsymbol{A}=\begin{bmatrix}1 & 0 & -1\\1 & -3 & 0\\4 & -13 & 1\end{bmatrix}$，$\lambda=-2$

15. $\boldsymbol{A}=\begin{bmatrix}4 & 2 & 3\\-1 & 1 & -3\\2 & 4 & 9\end{bmatrix}$，$\lambda=3$

16. $\boldsymbol{A}=\begin{bmatrix}3 & 0 & 2 & 0\\1 & 3 & 1 & 0\\0 & 1 & 1 & 0\\0 & 0 & 0 & 4\end{bmatrix}$，$\lambda=4$

求习题17和18中矩阵的本征值.

17. $\begin{bmatrix}0 & 0 & 0\\0 & 2 & 5\\0 & 0 & -1\end{bmatrix}$

18. $\begin{bmatrix}4 & 0 & 0\\0 & 0 & 0\\1 & 0 & -3\end{bmatrix}$

19. 无须计算，试求 $\boldsymbol{A}=\begin{bmatrix}1 & 2 & 3\\1 & 2 & 3\\1 & 2 & 3\end{bmatrix}$的一个本征值，验证你的答案.

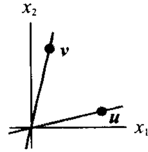
20. 无须计算，试求 $A=\begin{bmatrix}5&5&5\\5&5&5\\5&5&5\end{bmatrix}$ 的一个本征值和两个线性无关的本征向量，验证你的答案. 308

在习题 21 和 22 中，A 是一个 $n\times n$ 矩阵. 判断每个命题的真假，并说明理由.

21. a. 若 $Ax=\lambda x$ 对某个向量 x 成立，则 λ 是 A 的本征值.
    b. 矩阵 A 不可逆当且仅当 0 是 A 的本征值.
    c. 数量 c 是 A 的本征值当且仅当方程 $(A-cI)x=0$ 有非平凡解.
    d. 求 A 的本征向量可能很难，但判断一个给定的向量是否为 A 的本征向量却很容易.
    e. 为了求解 A 的本征值，需将 A 化简为阶梯形式.
22. a. 若 $Ax=\lambda x$ 对于某个数量 λ 成立，则 x 是 A 的本征向量.
    b. 若 v_1 和 v_2 是线性无关的本征向量，则它们对应于不同的本征值.
    c. 随机矩阵的一个稳态向量事实上是一个本征向量.
    d. 矩阵的本征值在它的主对角线上.
    e. A 的一个本征空间是某个矩阵的零空间.
23. 解释为什么一个 2×2 矩阵至多有两个不同的本征值. 解释为什么一个 $n\times n$ 矩阵至多有 n 个不同的本征值.
24. 构造一个有相同本征值的 2×2 矩阵.
25. 设 λ 为可逆矩阵 A 的一个本征值. 证明 λ^{-1} 为 A^{-1} 的一个本征值. [提示：假设非零向量 x 满足 $Ax=\lambda x$]
26. 证明：若 A^2 为零矩阵，则 A 有唯一本征值 0.
27. 证明：λ 为 A 的本征值当且仅当 λ 为 A^T 的本征值. [提示：考察 $A-\lambda I$ 和 $A^T-\lambda I$ 有什么联系.]
28. 利用习题 27 完成定理 1 中 A 是下三角矩阵时的证明.
29. 设 A 为 $n\times n$ 矩阵，且每一行元素的和都等于 s. 证明 s 为 A 的一个本征值. [提示：找一个本征向量.]
30. 设 A 为 $n\times n$ 矩阵，且每一列元素的和都等于 s. 证明 s 为 A 的一个本征值. [提示：利用习题 27 和 29.]

在习题 31 和 32 中，A 为线性变换 T 的矩阵，不必写出 A，求 A 的本征值并描述本征空间.

31. T 是 $\mathbf{R}^2$ 上以过原点的某条直线为对称轴的对称变换.
32. T 是 $\mathbf{R}^3$ 上以过原点的某条直线为旋转轴的旋转变换.
33. 设 u 和 v 是矩阵 A 的分别对应于本征值 λ 和 μ 的本征向量，并且设 c_1 和 c_2 为数量. 定义 $x_k=c_1\lambda^k u+c_2\mu^k v(k=0,1,2,\cdots)$
    a. 根据定义，如何表示 x_{k+1}？
    b. 根据 x_k 的公式计算 Ax_k，并证明 $Ax_k=x_{k+1}$. 这也证明了上述定义的序列 $\{x_k\}$ 满足差分方程 $x_{k+1}=Ax_k(k=0,1,2,\cdots)$.
34. 如果给定了一个初始向量 x_0，但 x_0 不是 A 的本征向量. 试说明如何构造差分方程 $x_{k+1}=Ax_k$ 的一个解. [提示：怎样将 x_0 与 A 的本征向量联系起来？]



35. u 和 v 为右图所示的向量，设 u 和 v 分别为 2×2 矩阵 A 的对应于本征值 2 和 3 的本征向量. $T:\mathbf{R}^2\to\mathbf{R}^2$ 为线性变换，且对一切 $x\in\mathbf{R}^2$，$T(x)=Ax$，令 $w=u+v$. 在图中画出 $T(u)$，$T(v)$ 和 $T(w)$.
36. 设 u 和 v 分别为矩阵 A 的对应于本征值 -1 和 3 的本征向量，重做习题 35.

[M]在习题 37～40 中，用矩阵软件求解矩阵的本征值，然后用例题 4 中行化简的方法求每个本征空间的一组基.

37. $\begin{bmatrix} 8 & -10 & -5 \\ 2 & 17 & 2 \\ -9 & -18 & 4 \end{bmatrix}$

309

38. $\begin{bmatrix} 9 & -4 & -2 & -4 \\ -56 & 32 & -28 & 44 \\ -14 & -14 & 6 & -14 \\ 42 & -33 & 21 & -45 \end{bmatrix}$

39. $\begin{bmatrix} 4 & -9 & -7 & 8 & 2 \\ -7 & -9 & 0 & 7 & 14 \\ 5 & 10 & 5 & -5 & -10 \\ -2 & 3 & 7 & 0 & 4 \\ -3 & -13 & -7 & 10 & 11 \end{bmatrix}$

40. $\begin{bmatrix} -4 & -4 & 20 & -8 & -1 \\ 14 & 12 & 46 & 18 & 2 \\ 6 & 4 & -18 & 8 & 1 \\ 11 & 7 & -37 & 17 & 2 \\ 18 & 12 & -60 & 24 & 5 \end{bmatrix}$

基础练习答案

1. 5 为 $\boldsymbol{A}$ 的本征向量当且仅当方程$(\boldsymbol{A}-5\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$ 有非平凡解.

$$\boldsymbol{A}-5\boldsymbol{I}=\begin{bmatrix} 6 & -3 & 1 \\ 3 & 0 & 5 \\ 2 & 2 & 6 \end{bmatrix}-\begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}=\begin{bmatrix} 1 & -3 & 1 \\ 3 & -5 & 5 \\ 2 & 2 & 1 \end{bmatrix}$$

对其增广矩阵进行行化简：

$$\begin{bmatrix} 1 & -3 & 1 & 0 \\ 3 & -5 & 5 & 0 \\ 2 & 2 & 1 & 0 \end{bmatrix}\sim\begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 0 & 8 & -1 & 0 \end{bmatrix}\sim\begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 0 & 0 & -5 & 0 \end{bmatrix}$$

易知齐次方程组没有自由变量. 从而 $\boldsymbol{A}-5\boldsymbol{I}$ 为可逆矩阵，也就是说 5 不是 $\boldsymbol{A}$ 的本征值.

2. 如果 $\boldsymbol{x}$ 是 $\boldsymbol{A}$ 的对应于本征值 λ 的本征向量，则 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$，从而

$$\boldsymbol{A}^2\boldsymbol{x}=\boldsymbol{A}(\lambda\boldsymbol{x})=\lambda\boldsymbol{Ax}=\lambda^2\boldsymbol{x}$$

再左乘 $\boldsymbol{A}$ 得，$\boldsymbol{A}^3\boldsymbol{x}=\boldsymbol{A}(\boldsymbol{A}^2\boldsymbol{x})=\boldsymbol{A}(\lambda^2\boldsymbol{x})=\lambda^2\boldsymbol{Ax}=\lambda^3\boldsymbol{x}$. 用归纳法可以证明一般形式为 $\boldsymbol{A}^k\boldsymbol{x}=\lambda^k\boldsymbol{x}$.

5.2 特征方程

方阵 $\boldsymbol{A}$ 的本征值的有用信息包含在一个特殊的数量方程中，这个方程称为 $\boldsymbol{A}$ 的特征方程. 我们先看一个简单的例子，然后再讨论一般情形.

【例题 1】 求 $\boldsymbol{A}=\begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$的本征值.

解：我们必须找出所有使矩阵方程

310

$$(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$$

有非平凡解的数量 λ. 根据 2.3 节可逆矩阵定理，这个问题等价于求所有使得矩阵 $\boldsymbol{A}-\lambda\boldsymbol{I}$不可逆的 λ，其中

$$\boldsymbol{A}-\lambda\boldsymbol{I}=\begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}-\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}=\begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix}$$

由 2.2 节定理 4 知，若行列式为零，则矩阵不可逆. 所以 $\boldsymbol{A}$ 的本征值为方程

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=\det\begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix}=0$$

的解. 又因为

$$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

所以

$$\begin{aligned}\det(\boldsymbol{A}-\lambda\boldsymbol{I}) &= (2-\lambda)(-6-\lambda)-(3)(3)\\ &= -12+6\lambda-2\lambda+\lambda^2-9\\ &= \lambda^2+4\lambda-21\end{aligned}$$

令 $\lambda^2+4\lambda-21=0$，得 $(\lambda-3)(\lambda+7)=0$；所以 $\boldsymbol{A}$ 的本征值为 3 和 -7. ■

例题 1 中的行列式将含有两个未知量(λ 和 $\boldsymbol{x}$)的矩阵方程 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$ 转化为只含一个未知量的数量方程 $\lambda^2+4\lambda-21=0$. 这一思想对 $n\times n$ 矩阵也有效. 尽管如此，在讨论大规模的矩阵以前，我们先总结一下研究本征值需用到的行列式的性质.

5.2.1 行列式

设 $\boldsymbol{A}$ 为 $n\times n$ 矩阵，$\boldsymbol{U}$ 为 $\boldsymbol{A}$ 通过行替换和行交换(没有数乘)而得到的阶梯形矩阵，并设 r 为所做行交换的次数. 则 $\boldsymbol{A}$ 的**行列式**(determinant)，记作 $\det \boldsymbol{A}$，为 $\boldsymbol{U}$ 中对角线元素 $u_{11},\cdots,u_{nn}$ 乘积的 $(-1)^r$ 倍. 如果 $\boldsymbol{A}$ 可逆，则 u_{11}，…，u_{nn} 都为主元(因为 $\boldsymbol{A}\sim\boldsymbol{I}_n$ 且 u_{ii} 没有被化为 1). 否则，至少 u_{nn} 为零，所以乘积 $u_{11}\cdots u_{nn}$ 为零. 这样就有[1]

$$\det \boldsymbol{A} = \begin{cases}(-1)^r\cdot(\boldsymbol{U}\text{ 中主元的乘积}), & \boldsymbol{A}\text{ 可逆}\\ 0, & \boldsymbol{A}\text{ 不可逆}\end{cases} \quad (1)$$

311

【例题 2】 设 $\boldsymbol{A}=\begin{bmatrix}1 & 5 & 0\\ 2 & 4 & -1\\ 0 & -2 & 0\end{bmatrix}$，计算 $\det \boldsymbol{A}$.

解：下面的行变换进行了一次行交换：

$$\boldsymbol{A}\sim\begin{bmatrix}1 & 5 & 0\\ 0 & -6 & -1\\ 0 & -2 & 0\end{bmatrix}\sim\begin{bmatrix}1 & 5 & 0\\ 0 & -2 & 0\\ 0 & -6 & -1\end{bmatrix}\sim\begin{bmatrix}1 & 5 & 0\\ 0 & -2 & 0\\ 0 & 0 & -1\end{bmatrix}=\boldsymbol{U}_1$$

所以 $\det \boldsymbol{A}=(-1)^1(1)(-2)(-1)=-2$. 下面所列的另一种变换没有使用行交换，它得到了另一个阶梯形矩阵. 最后一步是将第 2 行的 $-1/3$ 倍加到第 3 行：

$$\boldsymbol{A}\sim\begin{bmatrix}1 & 5 & 0\\ 0 & -6 & -1\\ 0 & -2 & 0\end{bmatrix}\sim\begin{bmatrix}1 & 5 & 0\\ 0 & -6 & -1\\ 0 & 0 & 1/3\end{bmatrix}=\boldsymbol{U}_2$$

此时 $\det \boldsymbol{A}=(-1)^0(1)(-6)(1/3)=-2$，与前面一致. ■

关于行列式的公式(1)表明 $\boldsymbol{A}$ 可逆当且仅当 $\det \boldsymbol{A}$ 非零. 这个事实和 5.1 节中发现的矩阵可逆的一种刻画将被加入可逆矩阵定理.

---

1. 公式(1)是在 3.2 节中推导出的. 没学过第三章的读者可以将这个公式作为 $\det \boldsymbol{A}$ 的定义. 一个重要的值得注意的事实是，对于 $\det \boldsymbol{A}$，若不使用数乘则从 $\boldsymbol{A}$ 得到的任何阶梯矩阵 $\boldsymbol{U}$ 都有相同的值.

【定理】　可逆矩阵定理(续)

设 $\boldsymbol{A}$ 为 $n\times n$ 矩阵，则 $\boldsymbol{A}$ 可逆当且仅当：

s. 数0不是 $\boldsymbol{A}$ 的本征值.

t. $\boldsymbol{A}$ 的行列式不为0.

当 $\boldsymbol{A}$ 为 3×3 矩阵，$|\det \boldsymbol{A}|$ 为 $\boldsymbol{A}$ 的列向量 $\boldsymbol{a}_1$，$\boldsymbol{a}_2$，$\boldsymbol{a}_3$ 所确定的平行六面体的体积,如图5-5所示(详见3.3节). 其体积非零当且仅当向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 线性无关且矩阵 $\boldsymbol{A}$ 可逆.(如果向量非零且线性相关,则它们位于同一平面或同一直线上.)

312

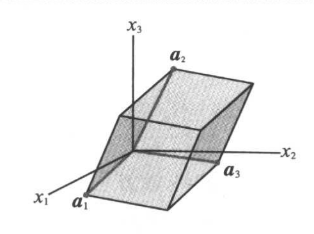


图5-5　$\boldsymbol{a}_1$，$\boldsymbol{a}_2$，$\boldsymbol{a}_3$ 确定的平行六面体

下述定理列出了本节要用到的3.1和3.2节中的一些事实. 为便于参考，(a)也列于其中.

【定理3】　行列式的性质

设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为 $n\times n$ 矩阵.

a. $\boldsymbol{A}$ 可逆当且仅当 $\det \boldsymbol{A}\neq 0$.

b. $\det \boldsymbol{AB}=(\det \boldsymbol{A})(\det \boldsymbol{B})$.

c. $\det \boldsymbol{A}^{\mathrm{T}}=\det \boldsymbol{A}$.

d. 若 $\boldsymbol{A}$ 为三角矩阵，则为 $\det \boldsymbol{A}$ 为 $\boldsymbol{A}$ 的主对角线元素的乘积.

e. 对 $\boldsymbol{A}$ 施行一次初等行变换：行替换不改变行列式的值；交换两行改变行列式的符号；数乘变换使行列式缩放相同的倍数.

5.2.2　特征方程

由定理3(a)，我们可以用行列式来判定矩阵 $\boldsymbol{A}-\lambda\boldsymbol{I}$ 何时不可逆. 数量方程 $\det(\boldsymbol{A}-\lambda\boldsymbol{I})=\boldsymbol{0}$ 称为 $\boldsymbol{A}$ 的**特征方程**(characteristic equation)，并且例题1中的说明证明了如下事实.

数量 λ 是 $n\times n$ 矩阵 $\boldsymbol{A}$ 的本征值当且仅当 λ 满足特征方程

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=0$$

【例题3】　求下列矩阵 $\boldsymbol{A}$ 的特征方程：

$$\boldsymbol{A}=\begin{bmatrix}5&-2&6&-1\\0&3&-8&0\\0&0&5&4\\0&0&0&1\end{bmatrix}$$

解： 写出 $\boldsymbol{A}-\lambda\boldsymbol{I}$，由定理3(d)：

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=\det\begin{bmatrix}5-\lambda&-2&6&-1\\0&3-\lambda&-8&0\\0&0&5-\lambda&4\\0&0&0&1-\lambda\end{bmatrix}=(5-\lambda)(3-\lambda)(5-\lambda)(1-\lambda)$$

特征方程为

313

$$(5-\lambda)^2(3-\lambda)(1-\lambda)=0$$

或

$$(\lambda-5)^2(\lambda-3)(\lambda-1)=0$$

展开乘积，我们也可以写成

$$\lambda^4-14\lambda^3+68\lambda^2-130\lambda+75=0$$ ■

在例题1和例题3中，$\det(\boldsymbol{A}-\lambda\boldsymbol{I})$是关于$\lambda$的一个多项式. 可以证明如果$\boldsymbol{A}$为$n\times n$矩阵，则$\det(\boldsymbol{A}-\lambda\boldsymbol{I})$是一个次数为$n$的多项式，这个多项式称为$\boldsymbol{A}$的**特征多项式**(characteristic polynomial).

例题3中，因为$(\lambda-5)$作为特征多项式的一个因式出现了两次，所以5称为重数为2的本征值. 一般而言，本征值λ的(代数)重数[(algebraic) multiplicity]就是它作为特征方程根的重数.

【例题4】 一个6×6矩阵的特征多项式是$\lambda^6-4\lambda^5-12\lambda^4$,求本征值及其重数.

解： 将多项式因式分解

$$\lambda^6-4\lambda^5-12\lambda^4=\lambda^4(\lambda^2-4\lambda-12)=\lambda^4(\lambda-6)(\lambda+2)$$

本征值为0(重数4)，6(重数1)和-2(重数1). ■

我们可以如下列出例题4中的本征值：0,0,0,0,6和-2，其中每个本征值重复出现的次数等于各自的重数.

因为$n\times n$矩阵的特征方程包含一个n次项，倘若允许复根，并将重数计算在内，则特征方程有n个根. 这样的复根称为复本征值，将在5.5节中讨论. 目前我们只考虑实本征值，并且数量仍为实数.

特征方程有着十分重要的理论意义. 在实际工作中，任何一个大于2×2的矩阵的本征值都应该用计算机求解得出，除非矩阵为三角形矩阵或具有其他特殊形式. 尽管动手计算一个3×3矩阵的特征多项式是容易的，但对其分解因式却比较困难(除非矩阵是特意挑选的). 参看本节最后的注记.

5.2.3 相似

下面的定理将阐述特征多项式的一个用法，并且为几个逼近本征值的迭代方法奠定了基础. 设$\boldsymbol{A}$和$\boldsymbol{B}$是$\boldsymbol{n\times n}$矩阵，如果存在一个可逆矩阵$\boldsymbol{P}$使得$\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{B}$，或等价地$\boldsymbol{A}=\boldsymbol{PBP}^{-1}$，则称$\boldsymbol{A}$**相似于**$\boldsymbol{B}$($\boldsymbol{A}$ is similar to $\boldsymbol{B}$). 令$\boldsymbol{Q}$为$\boldsymbol{P}^{-1}$，则$\boldsymbol{Q}^{-1}\boldsymbol{BQ}=\boldsymbol{A}$. 所以$\boldsymbol{B}$也相似于$\boldsymbol{A}$，并且我们简称为$\boldsymbol{A}$和$\boldsymbol{B}$**相似**(similar). 将$\boldsymbol{A}$转换为$\boldsymbol{P}^{-1}\boldsymbol{AP}$称为**相似变换**(similarity transformation).

314

【定理4】 **如果$n\times n$矩阵A和B相似，则它们有相同的特征多项式，从而有相同的本征值(具有相同的重数).**

证明： 如果$\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{AP}$，则

$$\boldsymbol{B}-\lambda\boldsymbol{I}=\boldsymbol{P}^{-1}\boldsymbol{AP}-\lambda\boldsymbol{P}^{-1}\boldsymbol{P}=\boldsymbol{P}^{-1}(\boldsymbol{AP}-\lambda\boldsymbol{P})=\boldsymbol{P}^{-1}(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{P}$$

利用定理3(b)的乘积性质，计算

$$\det(\boldsymbol{B}-\lambda\boldsymbol{I})=\det[\boldsymbol{P}^{-1}(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{P}]=\det(\boldsymbol{P}^{-1})\cdot\det(\boldsymbol{A}-\lambda\boldsymbol{I})\cdot\det(\boldsymbol{P}) \tag{2}$$

因为$\det(\boldsymbol{P}^{-1})\cdot\det(\boldsymbol{P})=\det(\boldsymbol{P}^{-1}\boldsymbol{P})=\det(\boldsymbol{I})=1$,从(2)我们可以得到

$$\det(\boldsymbol{B}-\lambda\boldsymbol{I})=\det(\boldsymbol{A}-\lambda\boldsymbol{I})$$ ■

警告:相似性与行等价不同(如果存在可逆矩阵$\boldsymbol{E}$使得$\boldsymbol{B}=\boldsymbol{EA}$,则称$\boldsymbol{A}$行等价于$\boldsymbol{B}$). 行变换通常会改变矩阵的本征值.

5.2.4　动态系统中的应用

本征值和本征向量是研究本章开篇实例介绍中提到的离散变化的动态系统的关键.

【例题5】　设 $A=\begin{bmatrix}0.95 & 0.03\\ 0.05 & 0.97\end{bmatrix}$. 动力系统定义为 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k(k=0,1,2\cdots)$，$\boldsymbol{x}_0=\begin{bmatrix}0.6\\ 0.4\end{bmatrix}$. 试分析该系统的长期行为.

解： 第一步是求 A 的本征值和每个本征空间的一组基. A 的特征方程为

$$0=\det\begin{bmatrix}0.95-\lambda & 0.03\\ 0.05 & 0.97-\lambda\end{bmatrix}=(0.95-\lambda)(0.97-\lambda)-(0.03)(0.05)$$
$$=\lambda^2-1.92\lambda+0.92$$

由二次公式

$$\lambda=\frac{1.92\pm\sqrt{(1.92)^2-4(0.92)}}{2}=\frac{1.92\pm\sqrt{0.0064}}{2}$$
$$=\frac{1.92\pm 0.08}{2}=1\quad 或\quad 0.92$$

315

容易验证,对应于 $\lambda=1$ 和 $\lambda=0.92$ 的本征向量分别是

$$\boldsymbol{v}_1=\begin{bmatrix}3\\ 5\end{bmatrix}和\ \boldsymbol{v}_2=\begin{bmatrix}1\\ -1\end{bmatrix}$$

的倍数.

下一步用 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 表示 $\boldsymbol{x}_0$. 这可以做到,因为 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 是 $\mathbf{R}^2$ 的一组基(为什么?). 所以存在系数 c_1 和 c_2 使得

$$\boldsymbol{x}_0=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2=[\boldsymbol{v}_1\quad \boldsymbol{v}_2]\begin{bmatrix}c_1\\ c_2\end{bmatrix}\tag{3}$$

事实上，

$$\begin{aligned}\begin{bmatrix}c_1\\ c_2\end{bmatrix}&=[\boldsymbol{v}_1\quad \boldsymbol{v}_2]^{-1}\boldsymbol{x}_0=\begin{bmatrix}3 & 1\\ 5 & -1\end{bmatrix}^{-1}\begin{bmatrix}0.60\\ 0.40\end{bmatrix}\\ &=\frac{1}{-8}\begin{bmatrix}-1 & -1\\ -5 & 3\end{bmatrix}\begin{bmatrix}0.60\\ 0.40\end{bmatrix}=\begin{bmatrix}0.125\\ 0.225\end{bmatrix}\end{aligned}\tag{4}$$

因为(3)中的 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 是 A 的本征向量，所以 $A\boldsymbol{v}_1=\boldsymbol{v}_1$ 且 $A\boldsymbol{v}_2=0.92\boldsymbol{v}_2$，我们容易算出每个 $\boldsymbol{x}_k$：

$$\begin{aligned}\boldsymbol{x}_1=A\boldsymbol{x}_0&=c_1A\boldsymbol{v}_1+c_2A\boldsymbol{v}_2 &&\text{根据 } \boldsymbol{x}\mapsto A\boldsymbol{x} \text{ 的线性性质}\\ &=c_1\boldsymbol{v}_1+c_2(0.92)\boldsymbol{v}_2 &&\boldsymbol{v}_1 \text{ 和 } \boldsymbol{v}_2 \text{ 是 } A \text{ 的本征向量}\\ \boldsymbol{x}_2=A\boldsymbol{x}_1&=c_1A\boldsymbol{v}_1+c_2(0.92)A\boldsymbol{v}_2\\ &=c_1\boldsymbol{v}_1+c_2(0.92)^2\boldsymbol{v}_2\end{aligned}$$

等等. 一般地

$$x_k = c_1 v_1 + c_2 (0.92)^k v_2 (k=0,1,2,\cdots)$$

代入(4)中的 c_1 和 c_2,

$$x_k = 0.125\begin{bmatrix}3\\5\end{bmatrix} + 0.225(0.92)^k\begin{bmatrix}1\\-1\end{bmatrix}(k=0,1,2,\cdots) \quad (5)$$

x_k 的这个显式公式给出了差分方程 $x_{k+1}=Ax_k$ 的解. 当 $k\to\infty$, $(0.92)^k$ 趋于 0 并且 x_k 趋于 $\begin{bmatrix}0.375\\0.625\end{bmatrix}=0.125v_1$. ■

例题 5 中的计算对 4.9 节中讨论的马尔可夫链有一个有趣的应用. 读过那节的读者可能会看出, 例题 5 中的矩阵 A 和 4.9 节中的迁移矩阵 M 是相同的, x_0 是城市与郊区人口分布的初始状态, x_k 表示 k 年之后的人口分布状态.

4.9 节定理 18 说明对于像 A 那样的矩阵, 序列 x_k 趋于一个稳态向量. 现在我们知道为什么 x_k 会这样. 稳态向量是 $0.125v_1$, 向量 v_1 的一个倍数, 并且 x_k 的计算公式(5)正好说明了为什么 $x_k\to 0.125v_1$.

316

注记

1. Mathematica 和 Maple 等计算机软件可以用符号计算求出中等规模矩阵的特征多项式. 但是对于 $n\geqslant 5$ 的 $n\times n$ 矩阵的特征方程, 并没有公式或有限算法可以求解.

2. 最好的求解本征向量的数值方法完全避免了特征多项式. 事实上, MATLAB 求矩阵 A 的特征多项式时, 首先求解 A 的本征值 $\lambda_1,\lambda_2,\cdots\lambda_n$, 然后将乘积 $(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$ 展开.

3. 估计矩阵 A 的本征值的几种常用算法都建立在定理 4 的基础上. 习题中会介绍十分有用的 QR 算法. 另一个称作*雅可比方法*的技术在 $A=A^T$ 时可以使用, 它计算形如

$$A_1 = A \text{ 和 } A_{k+1} = P_k^{-1}A_kP_k(k=1,2,\cdots)$$

的一列矩阵. 序列中每个矩阵都与 A 相似, 因而都和 A 有相同的本征值. 当 k 增加时, A_{k+1} 的非对角线元素趋于零, 对角线元素趋于 A 的本征值.

4. 另一些估计本征值的方法将在 5.8 节中讨论.

基础练习

求 $A=\begin{bmatrix}1&-4\\4&2\end{bmatrix}$ 的特征方程和本征值.

习题 5.2

求习题 1 ~ 8 中矩阵的特征多项式和本征值.

1. $\begin{bmatrix}2&7\\7&2\end{bmatrix}$　2. $\begin{bmatrix}5&3\\3&5\end{bmatrix}$　3. $\begin{bmatrix}3&-2\\1&-1\end{bmatrix}$　4. $\begin{bmatrix}5&-3\\-4&3\end{bmatrix}$

5. $\begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}$　6. $\begin{bmatrix} 3 & -4 \\ 4 & 8 \end{bmatrix}$　7. $\begin{bmatrix} 5 & 3 \\ -4 & 4 \end{bmatrix}$　8. $\begin{bmatrix} 7 & -2 \\ 2 & 3 \end{bmatrix}$

习题9～14需要3.1节中的技巧．用余子式展开或者用3.1节中习题15～18前介绍的用于计算3×3矩阵行列式的特定公式，求每个矩阵的特征多项式．［*注意*：因为含有参变量λ，所以只用行变换来求3×3矩阵的特征多项式是困难的．］

9. $\begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & -1 \\ 0 & 6 & 0 \end{bmatrix}$　10. $\begin{bmatrix} 0 & 3 & 1 \\ 3 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix}$　11. $\begin{bmatrix} 4 & 0 & 0 \\ 5 & 3 & 2 \\ -2 & 0 & 2 \end{bmatrix}$

12. $\begin{bmatrix} -1 & 0 & 1 \\ -3 & 4 & 1 \\ 0 & 0 & 2 \end{bmatrix}$　13. $\begin{bmatrix} 6 & -2 & 0 \\ -2 & 9 & 0 \\ 5 & 8 & 3 \end{bmatrix}$　14. $\begin{bmatrix} 5 & -2 & 3 \\ 0 & 1 & 0 \\ 6 & 7 & -2 \end{bmatrix}$

317

对于习题15～17中的矩阵，列出本征值，计重数．

15. $\begin{bmatrix} 4 & -7 & 0 & 2 \\ 0 & 3 & -4 & 6 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & 0 & 1 \end{bmatrix}$　16. $\begin{bmatrix} 5 & 0 & 0 & 0 \\ 8 & -4 & 0 & 0 \\ 0 & 7 & 1 & 0 \\ 1 & -5 & 2 & 1 \end{bmatrix}$　17. $\begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ -5 & 1 & 0 & 0 & 0 \\ 3 & 8 & 0 & 0 & 0 \\ 0 & -7 & 2 & 1 & 0 \\ -4 & 1 & 9 & -2 & 3 \end{bmatrix}$

18. 本征值λ的代数重数总是大于或等于λ所对应的本征空间的维数．求下列矩阵$\boldsymbol{A}$中的h，使得$\lambda=5$所对应的本征空间是2维的：

$$\boldsymbol{A}=\begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & h & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

19. 设$n\times n$矩阵$\boldsymbol{A}$有n个实本征值，计重数，设为$\lambda_1,\cdots,\lambda_n$，于是$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=(\lambda_1-\lambda)\cdot(\lambda_2-\lambda)\cdots(\lambda_n-\lambda)$．说明为什么$\det\boldsymbol{A}$是$\boldsymbol{A}$的本征值的乘积(考虑复特征根时，这个结论对任何方阵都成立)．

20. 用行列式的性质证明：$\boldsymbol{A}$和$\boldsymbol{A}^{\mathrm{T}}$有相同的特征多项式．

在习题21和22中，$\boldsymbol{A}$和$\boldsymbol{B}$是$n\times n$矩阵．判断每个命题的真假，并说明理由．

21. a. $\boldsymbol{A}$的行列式是其对角线元素的乘积．
　b. $\boldsymbol{A}$上的初等行变换不改变其行列式．
　c. $(\det\boldsymbol{A})(\det\boldsymbol{B})=\det\boldsymbol{AB}$
　d. 如果$\lambda+5$是$\boldsymbol{A}$的特征多项式的一个因式，则5是$\boldsymbol{A}$的本征值．

22. a. 如果$\boldsymbol{A}$为3×3矩阵，列向量为$\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$，则$\boldsymbol{A}$的行列式等于$\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$所确定的平行六面体的体积．
　b. $\det\boldsymbol{A}^{\mathrm{T}}=(-1)\det\boldsymbol{A}$
　c. $\boldsymbol{A}$的特征多项式的一个根r的重数，称为r作为$\boldsymbol{A}$的本征值的代数重数．
　d. $\boldsymbol{A}$上的行替换不改变它的本征值．

一个广泛使用的用于估计一般矩阵$\boldsymbol{A}$的本征值的方法是**QR**算法．在适当的条件下，这个算法产生一列矩阵，其中每一个都与$\boldsymbol{A}$相似，并且近似上三角矩阵，且对角线元素趋于$\boldsymbol{A}$的本征值．主要思想是将$\boldsymbol{A}$(或另一个与$\boldsymbol{A}$相似的矩阵)以$\boldsymbol{A}=\boldsymbol{Q}_1\boldsymbol{R}_1$的形式分解，其中$\boldsymbol{Q}_1^{\mathrm{T}}=\boldsymbol{Q}_1^{-1}$且$\boldsymbol{R}_1$为上三角矩阵．变换乘积顺序得到$\boldsymbol{A}_1=\boldsymbol{R}_1\boldsymbol{Q}_1$，并将其分解为$\boldsymbol{A}_1=\boldsymbol{Q}_2\boldsymbol{R}_2$；

然后得到 $A_2 = R_2Q_2$，等等. $A, A_1, \cdots$ 的相似性将由习题 23 中更一般的结论得到.

23. 证明：如果 $A = QR$，其中 Q 可逆，则 A 相似于 $A_1 = RQ$.

24. 证明：如果 A 和 B 相似，则 $\det A = \det B$.

25. 设 $A = \begin{bmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{bmatrix}, v_1 = \begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix}, x_0 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$. [注意：$A$ 是 4.9 节例题 5 中学习的随机矩阵.]

    a. 求 $\mathbf{R}^2$ 的由 v_1 和 A 的另一个本征向量 v_2 组成的基.

    b. 验证 x_0 可以写成 $x_0 = v_1 + cv_2$ 的形式.

    c. 对于 $k = 1, 2, \cdots$，定义 $x_k = A^k x_0$. 计算 x_1 和 x_2，并且写出 x_k 的计算公式. 然后证明 k 增加时 $x_k \to v_1$.

26. 设 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. 用关于行列式的公式(1)(在例题 2 前给出)证明 $\det A = ad - bc$. 考虑两种情况：$a \neq 0$ 和 $a = 0$.

27. 设 $A = \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.8 & 0.3 \\ 0.2 & 0 & 0.4 \end{bmatrix}$, $v_1 = \begin{bmatrix} 0.3 \\ 0.6 \\ 0.1 \end{bmatrix}$, $v_2 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$, $v_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ 和 $w = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

    a. 证明 v_1, v_2, v_3 是 A 的本征向量. [注意：A 是 4.9 节例题 3 中学习的随机矩阵.]

    b. 设 x_0 为 $\mathbf{R}^3$ 中任一向量，其元素非负且和为 1(在 4.9 节中，x_0 称为概率向量). 说明为什么存在常数 c_1, c_2, c_3 使得 $x_0 = c_1v_1 + c_2v_2 + c_3v_3$. 计算 $w^{\mathrm{T}}x_0$，并且推导出 $c_1 = 1$.

    c. 对于 $k = 1, 2, \cdots$，定义 $x_k = A^k x_0$，x_0 与(b)中相同. 证明 k 增加时 $x_k \to v_1$.

28. [M]构造一个 4×4 随机整数矩阵 A，验证 A 和 A^{T} 有相同的特征多项式(相同本征值的重数也相同). A 和 A^{T} 有相同的本征向量吗？对 5×5 矩阵做相同的分析. 报告你找到的矩阵和得到的结论.

318

29. [M]构造一个 4×4 随机整数矩阵 A.

    a. 不用行数乘，将 A 化简为阶梯矩阵 U，并且用公式(1)(在例题 2 前面)计算 $\det A$.（如果 A 恰好为奇异的，另选一随机矩阵重新计算.）

    b. 计算 A 的本征值和这些本征值的乘积(尽可能的精确).

    c. 列出矩阵 A，并且列出 A 的本征值和 U 中的主对角元，保留四位小数. 用你的矩阵软件计算 $\det A$，并与你在(a)和(b)中求出的结果作比较.

30. [M]设 $A = \begin{bmatrix} -6 & 28 & 21 \\ 4 & -15 & -12 \\ -8 & a & 25 \end{bmatrix}$. 对集合 $\{32, 31.9, 31.8, 32.1, 32.2\}$ 中的每个值 a，计算 A 的特征多项式和本征值. 对每种情形，画出特征多项式 $p(t) = \det(A - tI)$ 在区间 $0 \leq t \leq 3$ 上的图像. 如果可能，在一个坐标系中画出所有的图像. 试用图像说明本征值是如何随着 a 的变化而变化的.

基础练习答案

特征方程为

$$0 = \det(A - \lambda I) = \det\begin{bmatrix} 1-\lambda & -4 \\ 4 & 2-\lambda \end{bmatrix} = (1-\lambda)(2-\lambda) - (-4)(4) = \lambda^2 - 3\lambda + 18$$

由二次公式得，

$$\lambda=\frac{3\pm\sqrt{(-3)^2-4(18)}}{2}=\frac{3\pm\sqrt{-63}}{2}$$

显然，特征方程没有实数解，所以 A 没有实本征向量. 矩阵 A 作用在实向量空间 $\mathbf{R}^2$ 上，从而在 $\mathbf{R}^2$ 中不存在非零实向量 v，使得对于某个数量 λ 有 $Av=\lambda v$ 成立.

5.3 对 角 化

在许多情况中，关于本征值 - 本征向量的信息包含在一个矩阵中，该矩阵有 $A=PDP^{-1}$ 形式的分解. 本节中，这种分解使得对于较大的 k 值我们能够快速地计算 A^k，这也是线性代数很多应用中的一个基本思想. 稍后，在5.6和5.7节中，这种分解还将用于分析(和解耦)动态系统.

分解式中的 D 表示对角矩阵. 对于这样的 D，计算方幂十分容易.

【例题1】　若 $D=\begin{bmatrix}5&0\\0&3\end{bmatrix}$，则 $D^2=\begin{bmatrix}5&0\\0&3\end{bmatrix}\begin{bmatrix}5&0\\0&3\end{bmatrix}=\begin{bmatrix}5^2&0\\0&3^2\end{bmatrix}$ 且

$$D^3=DD^2=\begin{bmatrix}5&0\\0&3\end{bmatrix}\begin{bmatrix}5^2&0\\0&3^2\end{bmatrix}=\begin{bmatrix}5^3&0\\0&3^3\end{bmatrix}$$

一般地，

319

$$D^k=\begin{bmatrix}5^k&0\\0&3^k\end{bmatrix},\quad k\geqslant 1$$ ■

如果 $A=PDP^{-1}$，其中 P 可逆，D 为对角矩阵，则如下例所示，A^k 的计算也很容易.

【例题2】　设 $A=\begin{bmatrix}7&2\\-4&1\end{bmatrix}$，试求 A^k 的表达式，已知 $A=PDP^{-1}$，其中

$$P=\begin{bmatrix}1&1\\-1&-2\end{bmatrix} \text{且} D=\begin{bmatrix}5&0\\0&3\end{bmatrix}$$

解： 由 2×2 矩阵逆矩阵的标准公式得

$$P^{-1}=\begin{bmatrix}2&1\\-1&-1\end{bmatrix}$$

于是，由矩阵乘法结合律

$$A^2=(PDP^{-1})(PDP^{-1})=PD\underbrace{(P^{-1}P)}_{I}DP^{-1}=PDDP^{-1}$$

$$=PD^2P^{-1}=\begin{bmatrix}1&1\\-1&-2\end{bmatrix}\begin{bmatrix}5^2&0\\0&3^2\end{bmatrix}\begin{bmatrix}2&1\\-1&-1\end{bmatrix}$$

同理

$$A^3=(PDP^{-1})A^2=(PD\underbrace{P^{-1})P}_{I}D^2P^{-1}=PDD^2P^{-1}=PD^3P^{-1}$$

一般地，对于 $k\geqslant1$，

$$A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 2.5^k - 3^k & 5^k - 3^k \\ 2.3^k - 2.5^k & 2.3^k - 5^k \end{bmatrix}$$ ■

方阵 A 称作**可对角化**(diagonalizable)，如果 A 相似于对角矩阵，也就是说，如果对某个可逆矩阵 P 和某个对角矩阵 D 有 $A = PDP^{-1}$. 下述定理给出了可对角化矩阵的一个刻画并且说明了如何构造合适的分解.

【定理 5】 可对角化定理

$n \times n$ 矩阵 A 可对角化当且仅当 A 有 n 个线性无关的本征向量.

事实上，$A = PDP^{-1}$，其中 D 为对角矩阵，当且仅当 P 的列向量是 A 的 n 个线性无关的本征向量. 此时，D 的对角线元素是 A 的本征值，并且它们分别对应于 P 的本征向量.

320

换言之，A 可对角化当且仅当 A 有足够的本征向量可以构成 $\mathbf{R}^n$ 的一组基. 这样的一组基称为**本征向量基**(eigenvector basis).

证明： 首先，观察到如果 P 为任意 $n \times n$ 矩阵，其列向量为 $v_1, \cdots, v_n$，并且如果 D 为任意对角矩阵，其对角元为 $\lambda_1, \cdots, \lambda_n$，则

$$AP = A[v_1 \quad v_2 \quad \cdots \quad v_n] = [Av_1 \quad Av_2 \quad \cdots \quad Av_n] \tag{1}$$

并且

$$PD = P\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = [\lambda_1 v_1 \quad \lambda_2 v_2 \quad \cdots \quad \lambda_n v_n] \tag{2}$$

假设 A 可对角化且 $A = PDP^{-1}$. 两边同时右乘 P 得 $AP = PD$. 此时，由(1)，(2)得

$$[Av_1 \quad Av_2 \quad \cdots \quad Av_n] = [\lambda_1 v_1 \quad \lambda_2 v_2 \quad \cdots, \quad \lambda_n v_n] \tag{3}$$

由对应列相等可得：

$$Av_1 = \lambda_1 v_1, Av_2 = \lambda_2 v_2, \cdots, Av_n = \lambda_n v_n \tag{4}$$

因为 P 可逆，它的列向量一定线性无关. 此外，因为这些列向量非零，(4)说明 $\lambda_1, \cdots, \lambda_n$ 为本征值且 $v_1, \cdots, v_n$ 为对应的本征向量. 这个论述证明了定理的“必要性”.

最后，任给 n 个本征向量 $v_1, \cdots, v_n$，以它们作为 P 的列并且用对应的本征值 $\lambda_1, \cdots, \lambda_n$ 构造 D. 根据式(1)－式(3)，$AP = PD$. 这个事实对任意本征向量都成立. 实际上，如果本征向量线性无关，则 P 可逆(根据可逆矩阵定理)，并且由 $AP = PD$ 得到 $A = PDP^{-1}$. ■

5.3.1 矩阵的对角化

【例题 3】 如果可能，将下列矩阵对角化.

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

即，求可逆矩阵 P 和对角矩阵 D 使得 $A = PDP^{-1}$.

解： 需要四个步骤才能计算出定理 5 所要求的本征向量.

步骤1：求A的本征值．如5.2节所述，当矩阵大于2×2时，我们可以利用计算机求

321

矩阵的本征值．为了避免不必要的麻烦，本书通常会提供这个步骤所需要的信息．本例题中，特征方程是一个三次多项式，可以分解为：

$$0=\det(A-\lambda I)=-\lambda^3-3\lambda^2+4=-(\lambda-1)(\lambda+2)^2$$

本征值为$\lambda=1$和$\lambda=-2$．

步骤2：求A的三个线性无关的本征向量．需要三个向量是因为A是3×3矩阵．这是关键的一步．如果失败，则由定理5知A不可对角化．用5.1节中的方法，对每个本征空间都可以找到一组基：

$$\lambda=1\text{ 对应的基：}\boldsymbol{v}_1=\begin{bmatrix}1\\-1\\1\end{bmatrix}$$

$$\lambda=-2\text{ 对应的基：}\boldsymbol{v}_2=\begin{bmatrix}-1\\1\\0\end{bmatrix}\text{和 }\boldsymbol{v}_3=\begin{bmatrix}-1\\0\\1\end{bmatrix}$$

可以验证$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$为线性无关集．

步骤3：用步骤2中的向量构造P．向量的次序并不重要．按照步骤2中选定的次序，得

$$P=[\boldsymbol{v}_1\quad \boldsymbol{v}_2\quad \boldsymbol{v}_3]=\begin{bmatrix}1&-1&-1\\-1&1&0\\1&0&1\end{bmatrix}$$

步骤4：用相应的本征值构造D．在这个步骤中，关键是本征值的次序要与P中列向量选取的次序一致．对应于$\lambda=-2$的每个本征向量各用了一次，故需用本征值$\lambda=-2$两次：

$$D=\begin{bmatrix}1&0&0\\0&-2&0\\0&0&-2\end{bmatrix}$$

下面检验P，D为所求．为了避免计算P^{-1}，只需说明$AP=PD$．当P可逆时这与$A=PDP^{-1}$等价(故需确定P是可逆的!)．计算

$$AP=\begin{bmatrix}1&3&3\\-3&-5&-3\\3&3&1\end{bmatrix}\begin{bmatrix}1&-1&-1\\-1&1&0\\1&0&1\end{bmatrix}=\begin{bmatrix}1&2&2\\-1&-2&0\\1&0&-2\end{bmatrix}$$

322

$$PD=\begin{bmatrix}1&-1&-1\\-1&1&0\\1&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&-2&0\\0&0&-2\end{bmatrix}=\begin{bmatrix}1&2&2\\-1&-2&0\\1&0&-2\end{bmatrix}$$ ■

【例题4】　如果可能，将下列矩阵对角化．

$$A=\begin{bmatrix}2&4&3\\-4&-6&-3\\3&3&1\end{bmatrix}$$

解：$\boldsymbol{A}$ 的特征方程恰好和例题3中的一样：

$$0=\det(\boldsymbol{A}-\lambda\boldsymbol{I})=-\lambda^3-3\lambda^2+4=-(\lambda-1)(\lambda+2)^2$$

本征值为 $\lambda=1$ 和 $\lambda=-2$. 但当求解本征向量时，我们发现每个本征空间都是1维的.

$$\lambda=1\text{ 的基：}\boldsymbol{v}_1=\begin{bmatrix}1\\-1\\1\end{bmatrix}$$

$$\lambda=-2\text{ 的基：}\boldsymbol{v}_2=\begin{bmatrix}-1\\1\\0\end{bmatrix}$$

$\boldsymbol{A}$ 没有其他本征向量，并且每个 $\boldsymbol{A}$ 的本征向量都为 $\boldsymbol{v}_1$ 或 $\boldsymbol{v}_2$ 倍数. 因此 $\boldsymbol{A}$ 的本征向量不能构成 $\mathbf{R}^3$ 的一组基. 根据定理5，$\boldsymbol{A}$ 不可对角化. ■

【定理6】 有 n 个不同本征值的 $n\times n$ 矩阵可对角化.

证明：设 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$ 为 $\boldsymbol{A}$ 的对应于 n 个不同本征值的本征向量. 则由5.1节中定理2，$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$ 线性无关. 从而由定理5，$\boldsymbol{A}$ 可对角化. ■

$n\times n$ 矩阵有 n 个不同的本征值不是其可对角化的必要条件. 尽管例题3中的 3×3 矩阵只有两个不同的本征值，但它却可对角化.

【例题5】 判断下面的矩阵是否可对角化.

$$\boldsymbol{A}=\begin{bmatrix}5&-8&1\\0&0&7\\0&0&-2\end{bmatrix}$$

解：这很容易！因为该矩阵为三角矩阵，显然它的本征值为5，0和 -2. 从而 $\boldsymbol{A}$ 为有三个不同本征值的 3×3 矩阵，故 $\boldsymbol{A}$ 可对角化. ■ 323

5.3.2 有重复本征值的矩阵

如果 $n\times n$ 矩阵 $\boldsymbol{A}$ 有 n 个不同的本征值，对应的本征向量为 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$，并且如果 $\boldsymbol{P}=[\boldsymbol{v}_1\ \cdots\ \boldsymbol{v}_n]$，则由定理2，因为其列向量线性无关，$\boldsymbol{P}$ 显然可逆. 当 $\boldsymbol{A}$ 可对角化但有少于 n 个不同的本征值时，仍有方法可以构造 $\boldsymbol{P}$ 使得 $\boldsymbol{P}$ 可逆，如下面定理所述.[1]

【定理7】 设 $\boldsymbol{A}$ 为 $n\times n$ 矩阵，$\lambda_1,\cdots,\lambda_p$ 是 $\boldsymbol{A}$ 的互不相同的本征值.

a. 对于 $1\leqslant k\leqslant p$，λ_k 的本征空间的维数小于或等于本征值 λ_k 的重数.

b. 矩阵 $\boldsymbol{A}$ 可对角化当且仅当不同本征空间的维数之和等于 n，当且仅当每个 λ_k 所对应的本征空间的维数等于 λ_k 的重数.

c. 如果 $\boldsymbol{A}$ 可对角化并且对于每个 k，$\mathcal{B}_k$ 为对应于 λ_k 的本征空间的一组基，则 $\mathcal{B}_1,\cdots,\mathcal{B}_p$ 中的全体向量构成 $\mathbf{R}^n$ 的一组本征向量基.

---

1. 定理7的证明冗长但并不困难. 可参看 S. Friedberg, A. Insel, and L. Spence, *Linear Algebra*, 3rd ed. (Englewood Cliffs, NJ: Prentice-Hall, 1997), pp. 234 ~ 238.

【例题6】　如果可能，将下列矩阵对角化.

$$A=\begin{bmatrix}5&0&0&0\\0&5&0&0\\1&4&-3&0\\-1&-2&0&-3\end{bmatrix}$$

解：因为A为三角矩阵，所以本征值为5，-3，每一个的重数都是2. 用5.1节中的方法，求每个本征空间的一组基.

$$\lambda=5\text{ 对应的基：}v_1=\begin{bmatrix}-8\\4\\1\\0\end{bmatrix}\text{和 }v_2=\begin{bmatrix}-16\\4\\0\\1\end{bmatrix}$$

324

$$\lambda=-3\text{ 对应的基：}v_3=\begin{bmatrix}0\\0\\1\\0\end{bmatrix}\text{和 }v_4=\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

由定理7，$\{v_1,\cdots,v_4\}$线性无关. 所以矩阵$P=[v_1,\cdots,v_4]$可逆，并且$A=PDP^{-1}$，其中

$$P=\begin{bmatrix}-8&-16&0&0\\4&4&0&0\\1&0&1&0\\0&1&0&1\end{bmatrix}\text{和 }D=\begin{bmatrix}5&0&0&0\\0&5&0&0\\0&0&-3&0\\0&0&0&-3\end{bmatrix}$$ ■

基础练习

1. 计算A^8,其中$A=\begin{bmatrix}4&-3\\2&-1\end{bmatrix}$.
2. 设$A=\begin{bmatrix}-3&12\\-2&7\end{bmatrix}$, $v_1=\begin{bmatrix}3\\1\end{bmatrix}$并且$v_2=\begin{bmatrix}2\\1\end{bmatrix}$. 假定$v_1,v_2$为$A$的本征向量. 根据这些条件将$A$对角化.
3. 设A为4×4矩阵,本征值为5,3和-2,并且假定$\lambda=3$的本征空间是2维的. 根据这些条件是否能够判定A可对角化?

习题　5.3

在习题1和2中,设$A=PDP^{-1}$,计算A^4.

1. $P=\begin{bmatrix}5&7\\2&3\end{bmatrix}$, $D=\begin{bmatrix}2&0\\0&1\end{bmatrix}$

2. $P=\begin{bmatrix}2&-3\\-3&5\end{bmatrix}$, $D=\begin{bmatrix}1&0\\0&1/2\end{bmatrix}$

在习题3和4中,利用分解式$A=PDP^{-1}$计算A^k,其中k表示任意正整数.

3. $\begin{bmatrix}a&0\\3(a-b)&b\end{bmatrix}=\begin{bmatrix}1&0\\3&1\end{bmatrix}\begin{bmatrix}a&0\\0&b\end{bmatrix}\begin{bmatrix}1&0\\-3&1\end{bmatrix}$

4. $\begin{bmatrix}-2&12\\-1&5\end{bmatrix}=\begin{bmatrix}3&4\\1&1\end{bmatrix}\begin{bmatrix}2&0\\0&1\end{bmatrix}\begin{bmatrix}-1&4\\1&-3\end{bmatrix}$

在习题 5 和 6 中,矩阵 A 被分解为 $A=PDP^{-1}$ 的形式. 利用可对角化定理求解 A 的本征值以及每个本征空间的一组基.

5. $\begin{bmatrix}2&2&1\\1&3&1\\1&2&2\end{bmatrix}=\begin{bmatrix}1&1&2\\1&0&-1\\1&-1&0\end{bmatrix}\begin{bmatrix}5&0&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1/4&1/2&1/4\\1/4&1/2&-3/4\\1/4&-1/2&1/4\end{bmatrix}$

6. $\begin{bmatrix}4&0&-2\\2&5&4\\0&0&5\end{bmatrix}=\begin{bmatrix}-2&0&-1\\0&1&2\\1&0&0\end{bmatrix}\begin{bmatrix}5&0&0\\0&5&0\\0&0&4\end{bmatrix}\begin{bmatrix}0&0&1\\2&1&4\\-1&0&-2\end{bmatrix}$

如果可能,将习题 7 ~ 20 中的矩阵对角化. 习题 11 ~ 16 的本征值如下:(11)$\lambda=1, 2, 3$; (12)$\lambda=2, 8$; (13)$\lambda=5, 1$; (14)$\lambda=5, 4$; (15)$\lambda=3, 1$; (16)$\lambda=2, 1$. 对于习题 18, 已知一个本征值 $\lambda=5$, 一个本征向量为$(-2,1,2)$.

325

7. $\begin{bmatrix}1&0\\6&-1\end{bmatrix}$ 8. $\begin{bmatrix}5&1\\0&5\end{bmatrix}$ 9. $\begin{bmatrix}3&-1\\1&5\end{bmatrix}$ 10. $\begin{bmatrix}2&3\\4&1\end{bmatrix}$ 11. $\begin{bmatrix}-1&4&-2\\-3&4&0\\-3&1&3\end{bmatrix}$ 12. $\begin{bmatrix}4&2&2\\2&4&2\\2&2&4\end{bmatrix}$

13. $\begin{bmatrix}2&2&-1\\1&3&-1\\-1&-2&2\end{bmatrix}$ 14. $\begin{bmatrix}4&0&-2\\2&5&4\\0&0&5\end{bmatrix}$ 15. $\begin{bmatrix}7&4&16\\2&5&8\\-2&-1&-5\end{bmatrix}$ 16. $\begin{bmatrix}0&-4&-6\\-1&0&-3\\1&2&5\end{bmatrix}$

17. $\begin{bmatrix}4&0&0\\1&4&0\\0&0&5\end{bmatrix}$ 18. $\begin{bmatrix}-7&-16&4\\6&13&-2\\12&16&1\end{bmatrix}$ 19. $\begin{bmatrix}5&-3&0&9\\0&3&1&-2\\0&0&2&0\\0&0&0&2\end{bmatrix}$ 20. $\begin{bmatrix}4&0&0&0\\0&4&0&0\\0&0&2&0\\1&0&0&2\end{bmatrix}$

在习题 21 和 22 中, A, B, P 和 D 为 $n\times n$ 矩阵. 判断每个命题的真假, 并说明理由. (在做这些习题之前,认真复习定理 5、定理 6 和本节的例题.)

21. a. 如果对于某个矩阵 D 和可逆矩阵 P, $A=PDP^{-1}$, 则 A 可对角化.
    b. 如果 $\mathbf{R}^n$有由 A 的本征向量构成的一组基, 则 A 可对角化.
    c. A 可对角化当且仅当 A 有 n 个本征值(计重数).
    d. 如果 A 可对角化, 则 A 可逆.
22. a. 如果 A 有 n 个本征向量, 则 A 可对角化.
    b. 如果 A 可对角化, 则 A 有 n 个不同的本征值.
    c. 如果 $AP=PD$, 其中 D 是对角矩阵, 则 P 的非零列向量一定是 A 的本征向量.
    d. 如果 A 可逆, 则 A 可对角化.
23. A 为 5×5 矩阵, 有两个本征值. 一个本征空间是三维的, 另一个是二维的. A 可对角化吗? 为什么?
24. A 为 3×3 矩阵, 有两个本征值. 每个本征空间都是一维的. A 可对角化吗? 为什么?
25. A 为 4×4 矩阵, 有三个本征值. 一个本征空间是一维的, 另外两个本征空间其中之一是二维的. A 是否能对角化? 说明理由.
26. A 为 7×7 矩阵, 有三个本征值. 一个本征空间是二维的, 另外两个本征空间其中之一是三维的. A 是否能对角化? 说明理由.
27. 证明: 如果 A 既可对角化又可逆, 则 A^{-1}亦然.
28. 证明: 如果 A 有 n 个线性无关的本征向量, 则 A^{T} 亦然. [提示:使用可对角化定理.]
29. 分解式 $A=PDP^{-1}$不唯一. 用例题 2 中的矩阵 A 来说明这点. 令 $D_1=\begin{bmatrix}3&0\\0&5\end{bmatrix}$, 根据例题 2

中的条件求解矩阵 P_1 使得 $A=P_1D_1P_1^{-1}$.

30. A，D 如例题2中所示，求与例题2中 P 不相等的可逆矩阵 P_2，使得 $A=P_2DP_2^{-1}$.
31. 构造一个非零的 2×2 矩阵，使之可逆但不可对角化.
32. 构造一个非对角 2×2 矩阵，使之可对角化但不可逆.

[M]将习题33～36中的矩阵对角化. 用矩阵软件中的本征值命令求本征值，然后仿照5.1节求本征空间的基.

33. $\begin{bmatrix}-6&4&0&9\\-3&0&1&6\\-1&-2&1&0\\-4&4&0&7\end{bmatrix}$

34. $\begin{bmatrix}0&13&8&4\\4&9&8&4\\8&6&12&8\\0&5&0&-4\end{bmatrix}$

35. $\begin{bmatrix}11&-6&4&-10&-4\\-3&5&-2&4&1\\-8&12&-3&12&4\\1&6&-2&3&-1\\8&-18&8&-14&-1\end{bmatrix}$

36. $\begin{bmatrix}4&4&2&3&-2\\0&1&-2&-2&2\\6&12&11&2&-4\\9&20&10&10&-6\\15&28&14&5&-3\end{bmatrix}$

326

基础练习答案

1. $\det(A-\lambda I)=\lambda^2-3\lambda+2=(\lambda-2)(\lambda-1)$. 本征值为2和1，并且相应的本征向量为 $v_1=\begin{bmatrix}3\\2\end{bmatrix}$和 $v_2=\begin{bmatrix}1\\1\end{bmatrix}$. 接下来构造

$$P=\begin{bmatrix}3&1\\2&1\end{bmatrix},D=\begin{bmatrix}2&0\\0&1\end{bmatrix}\text{和}P^{-1}=\begin{bmatrix}1&-1\\-2&3\end{bmatrix}$$

因为 $A=PDP^{-1}$，

$$A^8=PD^8P^{-1}=\begin{bmatrix}3&1\\2&1\end{bmatrix}\begin{bmatrix}2^8&0\\0&1^8\end{bmatrix}\begin{bmatrix}1&-1\\-2&3\end{bmatrix}=\begin{bmatrix}3&1\\2&1\end{bmatrix}\begin{bmatrix}256&0\\0&1\end{bmatrix}\begin{bmatrix}1&-1\\-2&3\end{bmatrix}=\begin{bmatrix}766&-765\\510&-509\end{bmatrix}$$

2. 计算 $Av_1=\begin{bmatrix}-3&12\\-2&7\end{bmatrix}\begin{bmatrix}3\\1\end{bmatrix}=\begin{bmatrix}3\\1\end{bmatrix}=1\cdot v_1$，并且

$$Av_2=\begin{bmatrix}-3&12\\-2&7\end{bmatrix}\begin{bmatrix}2\\1\end{bmatrix}=\begin{bmatrix}6\\3\end{bmatrix}=3\cdot v_2$$

所以，v_1,v_2 分别为本征值1，3的本征向量. 因此

$$A=PDP^{-1}\text{，其中}P=\begin{bmatrix}3&2\\1&1\end{bmatrix},D=\begin{bmatrix}1&0\\0&3\end{bmatrix}$$

3. 是，A 可对角化. $\{v_1,v_2\}$为对应于 $\lambda=3$ 的本征空间的一组基. 另外，$\lambda=5$ 和 $\lambda=-2$ 分别至少有一本征向量，记作 v_3，v_4. 则$\{v_1,v_2,v_3,v_4\}$线性无关，根据定理7，A 可对角化. 不存在其他与 v_1,v_2,v_3,v_4 线性无关的本征向量，因为所有这些向量都属于 $\mathbf{R}^4$. 所以，$\lambda=5$ 和 $\lambda=-2$ 所对应的本征空间都是1维的.

5.4　本征向量和线性变换

本节的目的是从线性变换的角度来理解矩阵分解 $A=PDP^{-1}$. 当从适当的角度考

察变换 $\boldsymbol{x} \to A\boldsymbol{x}$ 时，它和简单的映射 $\boldsymbol{u} \to D\boldsymbol{u}$ 本质上是一样的. 甚至当 $\boldsymbol{D}$ 不是对角矩阵时，$\boldsymbol{A}$ 和 $\boldsymbol{D}$ 也有类似的联系.

回忆 1.9 节中，$\mathbf{R}^n$ 到 $\mathbf{R}^m$ 的任一线性变换 T 都可以通过左乘一个矩阵 $\boldsymbol{A}$ 来实现，该矩阵称作 T 的标准矩阵. 现在，我们要将这种表示推广到两个有限维空间上的任意线性变换.

327

5.4.1　线性变换的矩阵

设 V 为 n 维向量空间，W 为 m 维向量空间，且 T 为任一 V 到 W 的线性变换. 为了将 T 与矩阵联系起来，分别选定 V 和 W 的一组(有序)基 $\mathcal{B}$ 和 $\mathcal{C}$.

已知 V 中任一 $\boldsymbol{x}$，坐标向量 $[\boldsymbol{x}]_{\mathcal{B}} \in \mathbf{R}^n$，且它的像的坐标向量 $[T(\boldsymbol{x})]_{\mathcal{C}} \in \mathbf{R}^m$，如图 5-6 所示.

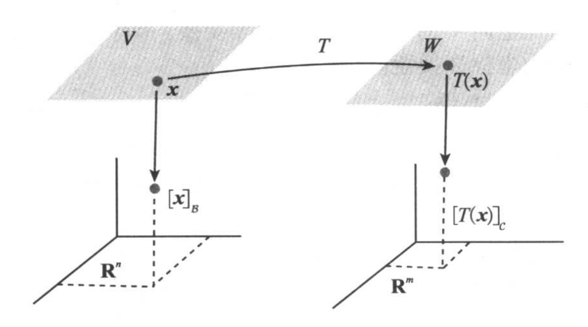


图 5-6　V 到 W 的线性变换

容易发现 $[\boldsymbol{x}]_{\mathcal{B}}$ 与 $[T(\boldsymbol{x})]_{\mathcal{C}}$ 之间的关系. 设 $\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$ 为 V 的基 $\mathcal{B}$. 若 $\boldsymbol{x} = r_1\boldsymbol{b}_1 + \cdots + r_n\boldsymbol{b}_n$，则

$$[\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}$$

且

$$T(\boldsymbol{x}) = T(r_1\boldsymbol{b}_1 + \cdots + r_n\boldsymbol{b}_n) = r_1T(\boldsymbol{b}_1) + \cdots + r_nT(\boldsymbol{b}_n) \tag{1}$$

因为 T 是线性的. 利用 W 中的基 $\mathcal{C}$，我们可以根据 $\mathcal{C}$ - 坐标向量来改写(1)：

$$[T(\boldsymbol{x})]_{\mathcal{C}} = r_1[T(\boldsymbol{b}_1)]_{\mathcal{C}} + \cdots + r_n[T(\boldsymbol{b}_n)_{\mathcal{C}}] \tag{2}$$

因为 $\mathcal{C}$ - 坐标向量属于 $\mathbf{R}^m$，向量方程(2)可以被写为矩阵方程，即：

$$[T(\boldsymbol{x})]_{\mathcal{C}} = \boldsymbol{M}[\boldsymbol{x}]_{\mathcal{B}} \tag{3}$$

其中

$$\boldsymbol{M} = [\,[T(\boldsymbol{b}_1)]_{\mathcal{C}} \quad [T(\boldsymbol{b}_2)]_{\mathcal{C}} \quad \cdots \quad [T(\boldsymbol{b}_n)]_{\mathcal{C}}] \tag{4}$$

矩阵 $\boldsymbol{M}$ 是 T 的一个矩阵表示,称为 **T 关于基 $\mathcal{B}$ 和 $\mathcal{C}$ 的矩阵**(matrix for T relative to the bases $\mathcal{B}$ and $\mathcal{C}$). 见图 5-7.

方程(3)说明，只要坐标向量是确定的，T 在 $\boldsymbol{x}$ 上的作用就可以看成左乘以 $\boldsymbol{M}$.

328

【例题 1】　设 $\mathcal{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ 为 V 的一组基，且 $\mathcal{C} = \{\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3\}$ 为 W 的一组基. 令 $T: V \to W$ 为具有如下性质的线性变换

$$T(\boldsymbol{b}_1)=3\boldsymbol{c}_1-2\boldsymbol{c}_2+5\boldsymbol{c}_3 \text{且 } T(\boldsymbol{b}_2)=4\boldsymbol{c}_1+7\boldsymbol{c}_2-\boldsymbol{c}_3$$

试求 T 关于基 $\mathcal{B}$ 和 $\mathcal{C}$ 的矩阵 M.

解： $\boldsymbol{b}_1$ 和 $\boldsymbol{b}_2$ 的像的 $\mathcal{C}$ - 坐标向量为

$$[T(\boldsymbol{b}_1)]_{\mathcal{C}}=\begin{bmatrix}3\\-2\\5\end{bmatrix} \quad \text{且} \quad [T(\boldsymbol{b}_2)]_{\mathcal{C}}=\begin{bmatrix}4\\7\\-1\end{bmatrix}$$

于是

$$M=\begin{bmatrix}3&4\\-2&7\\5&-1\end{bmatrix}$$

■

如果 $\mathcal{B}$ 和 $\mathcal{C}$ 为同一空间 V 的基，且 T 为恒等变换，即对任意 $\boldsymbol{x}\in V$，有 $T(\boldsymbol{x})=\boldsymbol{x}$，则(4)中的矩阵 $\boldsymbol{M}$ 恰为坐标变换矩阵(见 4.7 节).

5.4.2　V 到 V 上的线性变换

当 W 和 V 一样，且 $\mathcal{C}$ 和 $\mathcal{B}$ 一样时，(4)中的矩阵 $\boldsymbol{M}$ 称为 **T 关于基 $\mathcal{B}$ 的矩阵**(matrix for T relative to $\mathcal{B}$)，或简称为 **T 的 $\mathcal{B}$ - 矩阵**($\mathcal{B}$ - matrix for T)，记为 $[T]_{\mathcal{B}}$. 见图 5-8.

$T:V\to V$ 的 $\mathcal{B}$ - 矩阵满足

$$[T(\boldsymbol{x})]_{\mathcal{B}}=[T]_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}}, \text{对于 } V \text{ 中一切 } \boldsymbol{x} \tag{5}$$

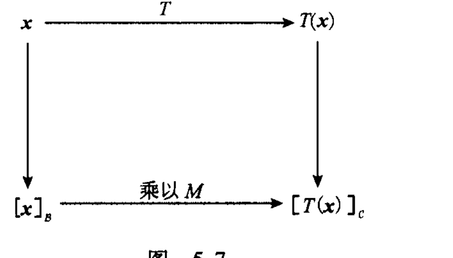


图 5-7

x　T　T(x)

[x]_B　乘以 [T]_B　[T(x)]_B

图 5-8

【例题 2】　定义为

$$T(a_0+a_1t+a_2t^2)=a_1+2a_2t$$

的映射 $T:\mathbf{P}_2\to\mathbf{P}_2$ 是一个线性变换.（学微积分的学生可以认出 T 为微分算子.）

a. 当 $\mathcal{B}$ 为基 $\{1,t,t^2\}$ 时，试求 T 的 $\mathcal{B}$ - 矩阵.

b. 对于 $\mathbf{P}_2$ 中的每个 $\boldsymbol{p}$ 验证 $[T(\boldsymbol{p})]_{\mathcal{B}}=[T]_{\mathcal{B}}[\boldsymbol{p}]_{\mathcal{B}}$.

解：

a. 计算基向量的像：

$$T(1)=0 \quad \text{零多项式}$$
$$T(t)=1 \quad \text{值恒为 1 的多项式}$$
$$T(t^2)=2t$$

329

然后写出 $T(1)$, $T(t)$ 和 $T(t^2)$ 的 $\mathcal{B}$ - 坐标向量(在本例题中通过观察可以得到)，并将它们放在一起得到 T 的 $\mathcal{B}$ - 矩阵：

$$[T(1)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad [T(t)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [T(t^2)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

$$[T]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

b. 对于一般的 $\boldsymbol{p}(t) = a_0 + a_1 t + a_2 t^2$,

$$[T(\boldsymbol{p})]_{\mathcal{B}} = [a_1 + 2a_2 t]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = [T]_{\mathcal{B}} [\boldsymbol{p}]_{\mathcal{B}}$$ ■

见图 5-9.

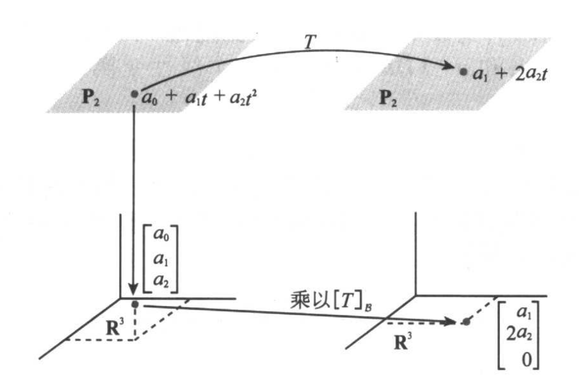


图 5-9 线性变换的矩阵表示

5.4.3 $\mathbf{R}^n$ 上的线性变换

在涉及 $\mathbf{R}^n$ 的应用问题中,线性变换绝大多数都是 $\boldsymbol{x} \mapsto A\boldsymbol{x}$ 形式的矩阵变换. 如果 A 可对角化, 则存在 $\mathbf{R}^n$ 的一组由 $\boldsymbol{A}$ 的本征向量构成的基 $\mathcal{B}$. 下面的定理 8 说明, 在这种情况下, $\boldsymbol{T}$ 的 $\mathcal{B}$ - 矩阵是对角矩阵. 将 $\boldsymbol{A}$ 对角化相当于求 $\boldsymbol{x} \mapsto A\boldsymbol{x}$ 的一个对角矩阵表示. 330

【定理 8】 对角矩阵表示

设 $\boldsymbol{A} = \boldsymbol{PDP}^{-1}$, 其中 $\boldsymbol{D}$ 为 $\boldsymbol{n} \times \boldsymbol{n}$ 对角矩阵. 如果 $\mathcal{B}$ 为由 $\boldsymbol{P}$ 的列向量构成的 $\mathbf{R}^n$ 的基, 则 $\boldsymbol{D}$ 为变换 $\boldsymbol{x} \mapsto A\boldsymbol{x}$ 的 $\mathcal{B}$ - 矩阵.

证明: 记 $\boldsymbol{P}$ 的列向量为 $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n$, 则 $\mathcal{B} = \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$ 且 $\boldsymbol{P} = [\boldsymbol{b}_1 \cdots \boldsymbol{b}_n]$. 此时, $\boldsymbol{P}$ 为 4.4 节中讨论的坐标变换矩阵 $\boldsymbol{P}_{\mathcal{B}}$, 其中

$$P[\boldsymbol{x}]_{\mathcal{B}} = \boldsymbol{x} \text{ 且 } [\boldsymbol{x}]_{\mathcal{B}} = \boldsymbol{P}^{-1}\boldsymbol{x}$$

若对 $\boldsymbol{x} \in \mathbf{R}^n$ 有 $T(\boldsymbol{x}) = A\boldsymbol{x}$, 则

$$\begin{aligned}[T]_{\mathcal{B}} &= [[T(\boldsymbol{b}_1)]_{\mathcal{B}}\cdots[T(\boldsymbol{b}_n)]_{\mathcal{B}}] && [T]_{\mathcal{B}}\text{ 的定义}\\ &= [[A\boldsymbol{b}_1]_{\mathcal{B}}\cdots[A\boldsymbol{b}_n]_{\mathcal{B}}] && \text{因为 } T(\boldsymbol{x})=A\boldsymbol{x}\\ &= [P^{-1}A\boldsymbol{b}_1\cdots P^{-1}A\boldsymbol{b}_n] && \text{坐标变换}\\ &= P^{-1}A[\boldsymbol{b}_1\cdots\boldsymbol{b}_n] && \text{矩阵乘法}\\ &= P^{-1}AP \end{aligned} \tag{6}$$

因为 $A=PDP^{-1}$，我们得到 $[T]_{\mathcal{B}}=P^{-1}AP=D$. ■

【例题3】　$T:\mathbf{R}^2\to\mathbf{R}^2$ 定义为 $T(\boldsymbol{x})=A\boldsymbol{x}$，其中 $A=\begin{bmatrix}7&2\\-4&1\end{bmatrix}$. 试求 $\mathbf{R}^2$ 的一组基 $\mathcal{B}$，使得 T 的 $\mathcal{B}$-矩阵为对角矩阵.

解：由5.3节中例题2知 $A=PDP^{-1}$，其中

$$P=\begin{bmatrix}1&1\\-1&-2\end{bmatrix}\text{且 } D=\begin{bmatrix}5&0\\0&3\end{bmatrix}$$

P 的列向量，记作 $\boldsymbol{b}_1$ 和 $\boldsymbol{b}_2$，是 A 的本征向量. 根据定理8，当 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 时 D 为 T 的 $\mathcal{B}$-矩阵. 映射 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 和 $\boldsymbol{u}\mapsto D\boldsymbol{u}$ 描述了同一个线性变换，只不过是相对于不同的基. ■

5.4.4　矩阵表示的相似

定理8的证明没有用到 D 是对角矩阵的条件. 因此，如果 A 相似于一个矩阵 C，其中 $A=PCP^{-1}$，则当基 $\mathcal{B}$ 由 P 的列向量构成时，C 是变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的 $\mathcal{B}$-矩阵. 矩阵分解 $A=PCP^{-1}$ 如图5-10所示.

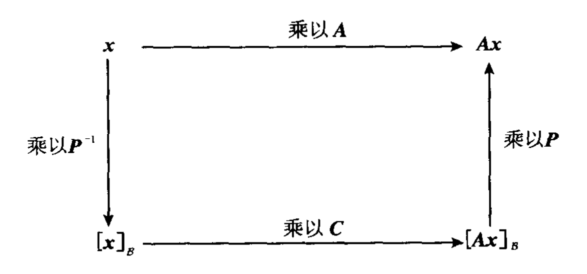


图5-10　两个矩阵表示的相似：$A=PCP^{-1}$

反过来，如果 $T:\mathbf{R}^n\to\mathbf{R}^n$ 定义为 $T(\boldsymbol{x})=A\boldsymbol{x}$，且如果 $\mathcal{B}$ 为 $\mathbf{R}^n$ 的任意一组基，则 T 的 $\mathcal{B}$-矩阵相似于 A. 事实上，(6)中的计算说明了如果矩阵 P 中的列都是 $\mathcal{B}$ 中的向量，则 $[T]_{\mathcal{B}}=P^{-1}AP$. 因此，相似于矩阵 A 的所有矩阵构成的集合与变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的所有矩阵表示构成的集合一致. 331

【例题4】　设 $A=\begin{bmatrix}4&-9\\4&8\end{bmatrix}$，$\boldsymbol{b}_1=\begin{bmatrix}3\\2\end{bmatrix}$，$\boldsymbol{b}_2=\begin{bmatrix}2\\1\end{bmatrix}$. A 的特征多项式为 $(\lambda+2)^2$，但本征值 -2 的本征空间只是一维的；所以 A 不可对角化. 尽管如此，基 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 仍具有这样的性质，即变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的 $\mathcal{B}$-矩阵是一个三角矩阵，该矩阵称为 A 的约当

形式.[1]试求这个 $\mathcal{B}$ – 矩阵.

解: 如果 $\boldsymbol{P}=[\boldsymbol{b}_1\quad \boldsymbol{b}_2]$，则 $\mathcal{B}$ – 矩阵为 $\boldsymbol{P}^{-1}\boldsymbol{AP}$. 计算

$$\boldsymbol{AP}=\begin{bmatrix}4 & -9\\ 4 & -8\end{bmatrix}\begin{bmatrix}3 & 2\\ 2 & 1\end{bmatrix}=\begin{bmatrix}-6 & -1\\ -4 & 0\end{bmatrix}$$

$$\boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{bmatrix}-1 & 2\\ 2 & -3\end{bmatrix}\begin{bmatrix}-6 & -1\\ -4 & 0\end{bmatrix}=\begin{bmatrix}-2 & 1\\ 0 & -2\end{bmatrix}$$

注意,$\boldsymbol{A}$ 的本征值在对角线上. ■

注记

计算 $\mathcal{B}$ – 矩阵 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 的一个有效方法是计算 $\boldsymbol{AP}$,然后将增广矩阵 $[\boldsymbol{P}\quad \boldsymbol{AP}]$ 行化简为 $[\boldsymbol{I}\quad \boldsymbol{P}^{-1}\boldsymbol{AP}]$. 单独计算 $\boldsymbol{P}^{-1}$ 是不必要的. 见 2.2 节习题 12.

基础练习

1. 如果 T 为 $\mathbf{P}_2$ 到 $\mathbf{P}_2$ 的线性变换，其关于 $\mathcal{B}=\{1,t,t^2\}$ 的矩阵为

$$[T]_{\mathcal{B}}=\begin{bmatrix}3 & 4 & 0\\ 0 & 5 & -1\\ 1 & -2 & 7\end{bmatrix}$$

试求 $T(a_0+a_1t+a_2t^2)$.

332

2. 设 $\boldsymbol{A}$，$\boldsymbol{B}$ 和 $\boldsymbol{C}$ 为 $n\times n$ 矩阵. 本书已经证明，如果 $\boldsymbol{A}$ 相似于 $\boldsymbol{B}$，则 $\boldsymbol{B}$ 相似于 $\boldsymbol{A}$. 这个性质和下面的命题一起说明了“相似”是一个等价关系(行等价是等价关系的另一个例子). 验证下面的 a. 和 b..

   a. $\boldsymbol{A}$ 相似于 $\boldsymbol{A}$.

   b. 如果 $\boldsymbol{A}$ 相似于 $\boldsymbol{B}$ 且 $\boldsymbol{B}$ 相似于 $\boldsymbol{C}$，则 $\boldsymbol{A}$ 相似于 $\boldsymbol{C}$.

习题 5.4

1. 设 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3\}$ 和 $\mathcal{D}=\{\boldsymbol{d}_1,\boldsymbol{d}_2\}$ 分别为向量空间 V 和 W 的基. 令 $T:V\to W$ 为具有下列性质的线性变换

$$T(\boldsymbol{b}_1)=3\boldsymbol{d}_1-5\boldsymbol{d}_2,T(\boldsymbol{b}_2)=-\boldsymbol{d}_1+6\boldsymbol{d}_2,T(\boldsymbol{b}_3)=4\boldsymbol{d}_2$$

试求 $\boldsymbol{T}$ 关于 $\mathcal{B}$ 和 D 的矩阵.

2. 设 $\mathcal{D}=\{\boldsymbol{d}_1,\boldsymbol{d}_2\}$ 和 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$ 分别为向量空间 V 和 W 的基. 令 T: $V\to W$ 为具有下列性质的线性变换

$$T(\boldsymbol{d}_1)=2\boldsymbol{b}_1-3\boldsymbol{b}_2,T(\boldsymbol{d}_2)=-4\boldsymbol{b}_1+5\boldsymbol{b}_2$$

试求 T 关于 $\mathcal{D}$ 和 $\mathcal{B}$ 的矩阵.

3. 设 $\varepsilon=\{\boldsymbol{e}_1,\ \boldsymbol{e}_2,\ \boldsymbol{e}_3\}$ 为 $\mathbf{R}^3$ 的标准基，$\mathcal{B}=\{\boldsymbol{b}_1,\ \boldsymbol{b}_2,\ \boldsymbol{b}_3\}$ 为向量空间 V 的一组基，且 $T:\mathbf{R}^3\to V$ 为具有下列性质的线性变换

$$T(x_1,x_2,x_3)=(x_3-x_2)\boldsymbol{b}_1-(x_1+x_3)\boldsymbol{b}_2+(x_1-x_2)\boldsymbol{b}_3$$

   a. 计算 $T(\boldsymbol{e}_1),T(\boldsymbol{e}_2)$ 和 $T(\boldsymbol{e}_3)$.

   b. 计算 $[T(\boldsymbol{e}_1)]_{\mathcal{B}}$，$[T(\boldsymbol{e}_2)]_{\mathcal{B}}$，和 $[T(\boldsymbol{e}_3)]_{\mathcal{B}}$.

   c. 试求 T 关于 ε 和 $\mathcal{B}$ 的矩阵.

---

1. 每个方阵 $\boldsymbol{A}$ 都相似于一个约当形矩阵. 用来产生若尔当形式的基由 $\boldsymbol{A}$ 的称为“生成本征向量”的本征向量组成. 见 *Applied Linear Algebra*, 3rd ed. (Englewood Cliffs, NJ: Prentice-Hall, 1988), by B. Noble and J. W. Daniel.

4. 设 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3\}$为向量空间 V 的一组基且 $T: V\to\mathbf{R}^2$ 为具有如下性质的线性变换

$$T(x_1\boldsymbol{b}_1+x_2\boldsymbol{b}_2+x_3\boldsymbol{b}_3)=\begin{bmatrix}2x_1-4x_2+5x_3\\ -x_2+3x_3\end{bmatrix}$$

试求 T 关于 $\mathcal{B}$ 的矩阵和 $\mathbf{R}^2$ 的标准基.

5. 设 T: $\mathbf{P}_2\to\mathbf{P}_3$ 为将多项式 $\boldsymbol{p}(t)$ 映为多项式 $(t+5)\boldsymbol{p}(t)$ 的变换.
   a. 试求 $\boldsymbol{p}(t)=2-t+t^2$ 的像.
   b. 证明 T 是线性变换.
   c. 试求 T 关于基 $\{1,t,t^2\}$ 和 $\{1,t,t^2,t^3\}$ 的矩阵.

6. 设 T: $\mathbf{P}_2\to\mathbf{P}_4$ 为将多项式 $\boldsymbol{p}(t)$ 映为多项式 $\boldsymbol{p}(t)+t^2\boldsymbol{p}(t)$ 的变换.
   a. 试求 $\boldsymbol{p}(t)=2-t+t^2$ 的像.
   b. 证明 T 是线性变换.
   c. 试求 T 关于基 $\{1,t,t^2\}$ 和 $\{1,t,t^2,t^3,t^4\}$ 的矩阵.

7. 设映射 T: $\mathbf{P}_2\to\mathbf{P}_2$ 定义为

$$T(a_0+a_1t+a_2t^2)=3a_0+(5a_0-2a_1)t+(4a_1+a_2)t^2$$

且 T 是线性的. 试求 T 关于基 $\mathcal{B}=\{1,t,t^2\}$ 的矩阵表示.

8. 设 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3\}$为向量空间 V 的一组基. T 为 V 到 V 的线性变换,其关于 $\mathcal{B}$ 的矩阵为

$$[T]_{\mathcal{B}}=\begin{bmatrix}0&-6&1\\0&5&-1\\1&-2&7\end{bmatrix}$$

试求 $T(3\boldsymbol{b}_1-4\boldsymbol{b}_2)$.

9. 定义 T: $\mathbf{P}_2\to\mathbf{R}^3$ 为 $T(\boldsymbol{p})=\begin{bmatrix}\boldsymbol{p}(-1)\\ \boldsymbol{p}(0)\\ \boldsymbol{p}(1)\end{bmatrix}$.
   a. 试求 $\boldsymbol{p}(t)=5+3t$ 在 T 之下的像.
   b. 证明 T 是线性变换.
   c. 试求 T 关于 $\mathbf{P}_2$ 的基 $\{1,t,t^2\}$ 和 $\mathbf{R}^3$ 的标准基的矩阵.

333

10. 定义 T: $\mathbf{P}_3\to\mathbf{R}^4$ 为 $T(\boldsymbol{p})=\begin{bmatrix}\boldsymbol{p}(-3)\\ \boldsymbol{p}(-1)\\ \boldsymbol{p}(1)\\ \boldsymbol{p}(3)\end{bmatrix}$.
   a. 证明 T 是线性变换.
   b. 试求 T 关于 $\mathbf{P}_3$ 的基 $\{1,t,t^2,t^3\}$ 和 $\mathbf{R}^4$ 的标准基的矩阵.

在习题 11 和 12 中，试求变换 $\boldsymbol{x}\to A\boldsymbol{x}$ 的 $\mathcal{B}$-矩阵，其中 $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$.

11. $A=\begin{bmatrix}3&4\\-1&-1\end{bmatrix}$, $\boldsymbol{b}_1=\begin{bmatrix}2\\-1\end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix}1\\2\end{bmatrix}$　12. $A=\begin{bmatrix}-1&4\\-2&3\end{bmatrix}$, $\boldsymbol{b}_1=\begin{bmatrix}3\\2\end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix}-1\\1\end{bmatrix}$

在习题 13～16 中，定义 T: $\mathbf{R}^2\to\mathbf{R}^2$ 为 $T(\boldsymbol{x})=A\boldsymbol{x}$. 试求 $\mathbf{R}^2$ 的一组基 $\mathcal{B}$，使得 $[T]_{\mathcal{B}}$ 为对角矩阵.

13. $A=\begin{bmatrix}0&1\\-3&4\end{bmatrix}$　14. $A=\begin{bmatrix}5&-3\\-7&1\end{bmatrix}$　15. $A=\begin{bmatrix}4&-2\\-1&3\end{bmatrix}$　16. $A=\begin{bmatrix}2&-6\\-1&3\end{bmatrix}$

17. 设 $A=\begin{bmatrix}1&1\\-1&3\end{bmatrix}$, $\mathcal{B}=\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$, 其中 $\boldsymbol{b}_1=\begin{bmatrix}1\\1\end{bmatrix}$, $\boldsymbol{b}_2=\begin{bmatrix}5\\4\end{bmatrix}$. 定义 T: $\mathbf{R}^2\to\mathbf{R}^2$ 为 $T(\boldsymbol{x})=A\boldsymbol{x}$.

a. 验证 $\boldsymbol{b}_1$ 为 $\boldsymbol{A}$ 的本征值但 $\boldsymbol{A}$ 不可对角化.

b. 试求 T 的 $\mathcal{B}$ - 矩阵.

18. 定义 T: $\mathbf{R}^3 \to \mathbf{R}^3$ 为 $T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$，其中 $\boldsymbol{A}$ 是本征值为 5 和 -2 的 3×3 矩阵. 是否存在 $\mathbf{R}^3$ 的一组基 $\mathcal{B}$ 使得 T 的 $\mathcal{B}$ - 矩阵为对角矩阵? 试加以讨论.

验证习题 19 ~ 24 中的陈述. 其中的矩阵为方阵.

19. 如果 $\boldsymbol{A}$ 可逆且相似于 $\boldsymbol{B}$，则 $\boldsymbol{B}$ 可逆且 $\boldsymbol{A}^{-1}$ 相似于 $\boldsymbol{B}^{-1}$. [提示:存在某个可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}$. 说明 $\boldsymbol{B}$ 可逆的原因,然后求一个可逆矩阵 $\boldsymbol{Q}$ 使得 $\boldsymbol{Q}^{-1}\boldsymbol{A}^{-1}\boldsymbol{Q} = \boldsymbol{B}^{-1}$.]

20. 如果 $\boldsymbol{A}$ 相似于 $\boldsymbol{B}$，则 $\boldsymbol{A}^2$ 相似于 $\boldsymbol{B}^2$.

21. 如果 $\boldsymbol{B}$ 相似于 $\boldsymbol{A}$，且 $\boldsymbol{C}$ 也相似于 $\boldsymbol{A}$，则 $\boldsymbol{B}$ 相似于 $\boldsymbol{C}$.

22. 如果 $\boldsymbol{A}$ 可对角化且 $\boldsymbol{B}$ 相似于 $\boldsymbol{A}$，则 $\boldsymbol{B}$ 亦可对角化.

23. 如果 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 且 $\boldsymbol{x}$ 为 $\boldsymbol{A}$ 的对应于本征值 λ 的本征向量，则 $\boldsymbol{P}^{-1}\boldsymbol{x}$ 亦为 $\boldsymbol{B}$ 的对应于 λ 的本征向量.

24. 如果 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似，则它们有相同的秩. [提示:参见第 4 章的补充题 13 和 14.]

25. 方阵 $\boldsymbol{A}$ 的迹是 $\boldsymbol{A}$ 对角线上的元素之和，记为 $\operatorname{tr} A$. 可以验证，对于任何两个 $n \times n$ 矩阵 $\boldsymbol{F}$ 和 $\boldsymbol{G}$，$\operatorname{tr}(\boldsymbol{F}\boldsymbol{G}) = \operatorname{tr}(\boldsymbol{G}\boldsymbol{F})$. 证明：如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相似，则 $\operatorname{tr} \boldsymbol{A} = \operatorname{tr} \boldsymbol{B}$.

26. 可以证明矩阵 $\boldsymbol{A}$ 的迹等于 $\boldsymbol{A}$ 的本征值之和. 当 $\boldsymbol{A}$ 可对角化时，验证这个事实.

27. 设 V 是基为 $\mathcal{B} = \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$的 $\mathbf{R}^n$; W 是具有标准基的 $\mathbf{R}^n$，记标准基为 ε; 考虑恒等变换 I: $\mathbf{R}^n \to \mathbf{R}^n$，其中 $I(\boldsymbol{x}) = \boldsymbol{x}$. 试求 I 关于 $\mathcal{B}$ 和 ε 的矩阵. 这个矩阵在 4.4 节中称作什么?

28. 设 V 是基为 $\mathcal{B} = \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$的一个向量空间; W 亦为空间 V 但基为 $\mathcal{C} = \{\boldsymbol{c}_1, \cdots, \boldsymbol{c}_n\}$，$I$ 为恒等变换 I: $V \to W$. 试求 I 关于 $\mathcal{B}$ 和 $\mathcal{C}$ 的矩阵. 这个矩阵在 4.7 节中称作什么?

29. 令 V 是基为 $\mathcal{B} = \{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n\}$的一个向量空间. 试求恒等变换 I: $V \to V$ 的 $\mathcal{B}$ - 矩阵.

[M] 在习题 30 和 31 中，当 $\mathcal{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3\}$时，试求变换 $\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$ 的 $\mathcal{B}$ - 矩阵.

30. $\boldsymbol{A} = \begin{bmatrix} -14 & 4 & -14 \\ -33 & 9 & -31 \\ 11 & -4 & 11 \end{bmatrix}$, $\boldsymbol{b}_1 = \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$, $\boldsymbol{b}_2 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$, $\boldsymbol{b}_3 = \begin{bmatrix} -1 \\ -2 \\ 0 \end{bmatrix}$

31. $\boldsymbol{A} = \begin{bmatrix} -7 & -48 & -16 \\ 1 & 14 & 6 \\ -3 & -45 & -19 \end{bmatrix}$, $\boldsymbol{b}_1 = \begin{bmatrix} -3 \\ 1 \\ -3 \end{bmatrix}$, $\boldsymbol{b}_2 = \begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix}$, $\boldsymbol{b}_3 = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}$

32. [M] 令 T 为标准矩阵如下的变换. 试求 $\mathbf{R}^4$ 的一组基,使得 $[T]_{\mathcal{B}}$ 为对角矩阵.

$$\boldsymbol{A} = \begin{bmatrix} 15 & -66 & -44 & -33 \\ 0 & 13 & 21 & -15 \\ 1 & -15 & -21 & 12 \\ 2 & -18 & -22 & 8 \end{bmatrix}$$

334

基础练习答案

1. 令 $\boldsymbol{p}(t) = a_0 + a_1 t + a_2 t^2$ 然后计算

$$[T(\boldsymbol{p})]_{\mathcal{B}} = [T]_{\mathcal{B}}[\boldsymbol{p}]_{\mathcal{B}} = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix}\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 3a_0 + 4a_1 \\ 5a_1 - a_2 \\ a_0 - 2a_1 + 7a_2 \end{bmatrix}$$

所以 $T(\boldsymbol{p}) = (3a_0 + 4a_1) + (5a_1 - a_2)t + (a_0 - 2a_1 + 7a_2)t^2$.

2. a. $\boldsymbol{A} = (\boldsymbol{I})^{-1}\boldsymbol{A}\boldsymbol{I}$，所以 $\boldsymbol{A}$ 相似于 $\boldsymbol{A}$.

b. 根据假设，存在可逆矩阵 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$ 使得 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 且 $\boldsymbol{C} = \boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q}$. 将 $\boldsymbol{B}$ 公式代入 $\boldsymbol{C}$ 的公式

中，且根据乘积的逆的一个性质得：

$$C=Q^{-1}BQ=Q^{-1}(P^{-1}AP)Q=(PQ)^{-1}A(PQ)$$

这个方程的形式恰好证明了 A 相似于 C.

5.5 复本征值

因为 $n\times n$ 矩阵的特征方程是一个次数为 n 的多项式，所以倘若包含复根，并且将重数计算在内，则该方程总是有 n 个根. 本节将说明如果实矩阵 A 的特征方程有某些复根，则这些根提供了关于 A 的重要的信息. 其中的关键是让 A 作用在复数 n 元组所构成的 $\mathbf{C}^n$ 空间上.[1]

我们对于 $\mathbf{C}^n$ 空间的兴趣并不是来自于要推广前面章节结论的愿望，尽管那在事实上可以开发线性代数的重要的新应用.[2] 恰当的说，本节关于复本征值的学习本质上是为了揭示那些实元素矩阵中"隐藏"的信息，而这些矩阵在很多现实问题中都会出现. 这样的问题包括许多实动态系统，如周期运动，摆动，或空间中某些类型的旋转.

已经得到的关于 $\mathbf{R}^n$ 的矩阵本征值 - 本征向量理论可以同样应用于 $\mathbf{C}^n$. 所以复数量 λ 满足 $\det(A-\lambda I)=0$ 当且仅当 $\mathbf{C}^n$ 中存在某个非零向量 x 使得 $Ax=\lambda x$. 称 λ 为**(复)本征值**((complex) eigenvalue)，x 为对应于 λ 的**(复)本征向量**((complex) eigenvector).

【例题1】　如果 $A=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$，则 $\mathbf{R}^2$ 上的线性变换 $x\mapsto Ax$ 将平面逆时针方向旋转了90度. A 的作用是周期的，因为在旋转四个90度后，向量回到了原来的位置. 显然，没有非零向量被映为它自己的倍数，所以 A 在 $\mathbf{R}^2$ 中没有本征向量，从而没有实本征值. 事实上，A 的特征方程为

335

$$\lambda^2+1=0$$

仅有的根为复数：$\lambda=\mathrm{i}$ 和 $\lambda=-\mathrm{i}$. 然而，如果我们将 A 作用在 $\mathbf{C}^2$ 上，则

$$\begin{bmatrix}0&-1\\1&0\end{bmatrix}\begin{bmatrix}1\\-\mathrm{i}\end{bmatrix}=\begin{bmatrix}\mathrm{i}\\1\end{bmatrix}=\mathrm{i}\begin{bmatrix}1\\-\mathrm{i}\end{bmatrix}$$

$$\begin{bmatrix}0&-1\\1&0\end{bmatrix}\begin{bmatrix}1\\\mathrm{i}\end{bmatrix}=\begin{bmatrix}-\mathrm{i}\\1\end{bmatrix}=-\mathrm{i}\begin{bmatrix}1\\\mathrm{i}\end{bmatrix}$$

于是 i 和 $-\mathrm{i}$ 为本征值，对应的本征向量为 $\begin{bmatrix}1\\-\mathrm{i}\end{bmatrix}$ 和 $\begin{bmatrix}1\\\mathrm{i}\end{bmatrix}$.（例题2中讨论了一种求复本征向量的方法.）■

本节的重点集中在下面这道例题的矩阵上.

---

1. 参考附录B中关于复数的简短讨论. 实向量空间上的矩阵代数和概念可以推广至有复元素和数量的情形. 特别地，$A(c\boldsymbol{x}+d\boldsymbol{y})=cA\boldsymbol{x}+dA\boldsymbol{y}$，其中 A 为复元素的 $m\times n$ 矩阵，$\boldsymbol{x},\boldsymbol{y}\in\mathbf{C}^n$，$c,d\in\mathbf{C}$.
2. 线性代数中级课程中经常讨论这个主题. 它们在电机工程中有着特别的重要性.

【例题 2】 设 $A=\begin{bmatrix}0.5 & -0.6\\ 0.75 & 1.1\end{bmatrix}$. 试求 A 的本征值以及每个本征空间的一组基.

解：A 的本征方程为

$$0=\det\begin{bmatrix}0.5-\lambda & -0.6\\ 0.75 & 1.1-\lambda\end{bmatrix}=(0.5-\lambda)(1.1-\lambda)-(-0.6)(0.75)=\lambda^2-1.6\lambda+1$$

由二次公式，$\lambda=\frac{1}{2}\left[1.6\pm\sqrt{(-1.6)^2-4}\right]=0.8\pm0.6i$. 对于本征值 $\lambda=0.8-0.6i$，构造

$$A-(0.8-0.6i)I=\begin{bmatrix}0.5 & -0.6\\ 0.75 & 1.1\end{bmatrix}-\begin{bmatrix}0.8-0.6i & 0\\ 0 & 0.8-0.6i\end{bmatrix}=\begin{bmatrix}-0.3+0.6i & -0.6\\ 0.75 & 0.3+0.6i\end{bmatrix} \quad (1)$$

由于涉及复数算术，用手算化简一般的增广矩阵相当麻烦. 不过，通过观察，我们可以大大简化计算：因为 $0.8-0.6i$ 是本征值，方程组

$$\begin{aligned}(-0.3+0.6i)x_1 - \quad 0.6x_2 &= 0\\ 0.75x_1+(0.3+0.6i)x_2&=0\end{aligned} \quad (2)$$

有非平凡解(其中 x_1 和 x_2 可能为复数). 因此，(2)中的两个方程确定了 x_1 和 x_2 之间的相同关系，并且每个方程都可以用一个变量来表示另一个变量.[1] 336

由(2)中第二个方程得

$$\begin{aligned}0.75x_1&=(-0.3-0.6i)x_2\\ x_1&=(-0.4-0.8i)x_2\end{aligned}$$

选取 $x_2=5$ 来消除小数，并且得到 $x_1=-2-4i$. 对应于 $\lambda=0.8-0.6i$ 的本征空间的一组基为

$$v_1=\begin{bmatrix}-2-4i\\ 5\end{bmatrix}$$

对 $\lambda=0.8+0.6i$ 进行类似的计算，可以得到本征向量

$$v_2=\begin{bmatrix}-2+4i\\ 5\end{bmatrix}$$

作为检验，计算

$$Av_2=\begin{bmatrix}0.5 & -0.6\\ 0.75 & 1.1\end{bmatrix}\begin{bmatrix}-2+4i\\ 5\end{bmatrix}=\begin{bmatrix}-4+2i\\ 4+3i\end{bmatrix}=(0.8+0.6i)v_2$$ ■

我们惊奇地发现，例题 2 中的矩阵 A 所确定的线性变换 $x\mapsto Ax$ 原来是一个旋转.

1. 得到这个结论的另一个方法是认识到(1)的矩阵不可逆，所以它的行向量线性相关(作为 $\mathbb{C}^2$ 中向量)，从而一个行向量是另一个的(复)倍数.

选取适当的点并绘出图像时，这一事实显而易见.

【例题3】　为了了解在例题2中乘以A的作用是如何影响点的，一种方法是画出任意一个初始点，比如说$x_0=(2,0)$，然后画出这个点反复乘以A时得到的像. 即画出

$$x_1=Ax_0=\begin{bmatrix}0.5 & -0.6\\0.75 & 1.1\end{bmatrix}\begin{bmatrix}2\\0\end{bmatrix}=\begin{bmatrix}1.0\\1.5\end{bmatrix}$$

$$x_2=Ax_1=\begin{bmatrix}0.5 & -0.6\\0.75 & 1.1\end{bmatrix}\begin{bmatrix}1.0\\1.5\end{bmatrix}=\begin{bmatrix}-0.4\\2.4\end{bmatrix}$$

$$x_3=Ax_2,\cdots$$

图5-11中稍大的点是$x_0,\cdots,x_8$，小的点则代表$x_9,\cdots,x_{100}$. 该点列沿着一个椭圆轨道分布. ■

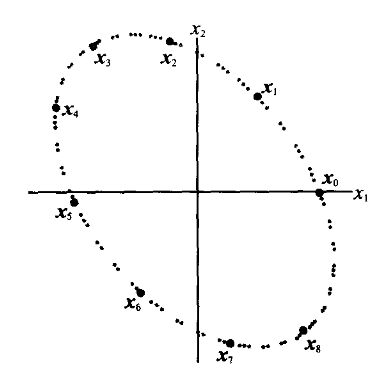


图5-11　点x_0在具有复本征值的矩阵作用下的迭代

当然，图5-11并没有解释为什么会发生旋转. 其中的奥秘隐藏于复本征向量的实部和虚部中.

5.5.1　向量的实部和虚部

$\mathbb{C}^n$中复向量x的复共轭是指$\mathbb{C}^n$中的向量$\bar{x}$，其元素为x中元素的复共轭. 复向量x的**实部**和**虚部**(real and imaginary parts)是指x中元素的实部和虚部所构成的向量

337

Re x 和 Im x.

【例题4】　如果$x=\begin{bmatrix}3-\mathrm{i}\\\mathrm{i}\\2+5\mathrm{i}\end{bmatrix}=\begin{bmatrix}3\\0\\2\end{bmatrix}+\mathrm{i}\begin{bmatrix}-1\\1\\5\end{bmatrix}$，则

$$\mathrm{Re}\,x=\begin{bmatrix}3\\0\\2\end{bmatrix},\ \mathrm{Im}\,x=\begin{bmatrix}-1\\1\\5\end{bmatrix},\ 且\ \bar{x}=\begin{bmatrix}3\\0\\2\end{bmatrix}-\mathrm{i}\begin{bmatrix}-1\\1\\5\end{bmatrix}=\begin{bmatrix}3+\mathrm{i}\\-\mathrm{i}\\2-5\mathrm{i}\end{bmatrix}$$ ■

如果B为$m\times n$矩阵，其中可能含有复元素，则记$\bar{B}$为元素是B中元素复共轭的矩阵. 复数共轭的性质同样适用于复矩阵代数：

$$\overline{rx}=\bar{r}\,\bar{x},\quad \overline{Bx}=\bar{B}\,\bar{x},\quad \overline{BC}=\bar{B}\,\bar{C},\quad 且\quad \overline{rB}=\bar{r}\,\bar{B}$$

5.5.2 作用在 $\mathbf{C}^n$ 上的实矩阵的本征值和本征向量

设 A 为 $n\times n$ 矩阵，其元素是实的，则 $\overline{Ax}=\overline{A}\,\overline{x}=A\overline{x}$. 如果 λ 是 A 的本征值，且 x 是 $\mathbf{C}^n$ 中对应的本征向量，则

$$A\overline{x}=\overline{Ax}=\overline{\lambda x}=\overline{\lambda}\,\overline{x}$$

于是 $\overline{\lambda}$ 也是 A 的本征值，其对应的本征向量是 $\overline{x}$. 这说明当 A 为实矩阵时，它的复本征值以共轭对的形式成对出现. 通常，我们所说的复本征值是指本征值 $\lambda=a+bi$，其中 $b\neq 0$.

【例题 5】 例题 2 中实矩阵的本征值是复共轭的，即 0.8 − 0.6i 和 0.8 + 0.6i. 例题 2 中求出的对应的本征向量也是共轭的： 338

$$v_1=\begin{bmatrix}-2-4i\\5\end{bmatrix} \text{且 } v_2=\begin{bmatrix}-2+4i\\5\end{bmatrix}=\overline{v}_1$$ ■

下面的例题为研究所有具有复本征值的实 2×2 矩阵奠定了基础.

【例题 6】 如果 $C=\begin{bmatrix}a & -b\\ b & a\end{bmatrix}$，其中 a 和 b 为实数且不全为零，则 C 的本征值为 $\lambda=a\pm bi$（见本节末尾的基础练习）. 此外，如果 $r=|\lambda|=\sqrt{a^2+b^2}$，则

$$C=r\begin{bmatrix}a/r & -b/r\\ b/r & a/r\end{bmatrix}=\begin{bmatrix}r & 0\\ 0 & r\end{bmatrix}\begin{bmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{bmatrix}$$

其中 φ 是 x 轴正向与以 $(0,0)$ 为端点经过 (a,b) 的射线之间的夹角. 见图 5-12 和附录 B. 角 φ 称为 $\lambda=a+bi$ 的辐角. 从而变换 $x\mapsto Cx$ 可以看成角度为 φ 的旋转和 $|\lambda|$ 数乘的复合（见图 5-13）. ■

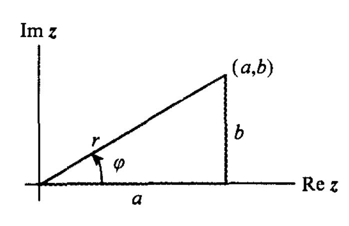


图 5-12 辐角 φ

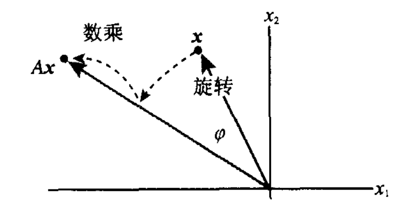


图 5-13 旋转和数乘的复合

最后，对于具有复本征值的实矩阵，我们即将揭示隐藏于其中的旋转.

【例题 7】 如例题 2，设 $A=\begin{bmatrix}0.5 & -0.6\\ 0.75 & 1.1\end{bmatrix}$，$\lambda=0.8-0.6i$ 且 $v_1=\begin{bmatrix}-2-4i\\5\end{bmatrix}$. 令 P 为 2×2 实矩阵

$$P=[\operatorname{Re} v_1 \quad \operatorname{Im} v_1]=\begin{bmatrix}-2 & -4\\ 5 & 0\end{bmatrix}$$

并且令

$$C=P^{-1}AP=\frac{1}{20}\begin{bmatrix}0 & 4\\ -5 & -2\end{bmatrix}\begin{bmatrix}0.5 & -0.6\\ 0.75 & 1.1\end{bmatrix}\begin{bmatrix}-2 & -4\\ 5 & 0\end{bmatrix}=\begin{bmatrix}0.8 & -0.6\\ 0.6 & 0.8\end{bmatrix}$$ 339

根据例题 6，因为 $|\lambda|^2=(0.8)^2+(0.6)^2=1$ 所以 C 只是一个旋转. 由 $C=P^{-1}AP$，得

$$A = PCP^{-1} = P\begin{bmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{bmatrix}P^{-1}$$

C 就是“隐藏”在 A 中的旋转！矩阵 P 给出了一个变量替换，即 $x = Pu$. A 的作用相当于先将 x 变量替换成 u，然后旋转，最后回到 Ax. 见图 5-14. 如图 5-11 所示，旋转产生了一个椭圆，而不是圆，因为由 P 的列向量所确定的坐标系不是正交的，并且两个坐标轴的单位长度不相等. ■

下面的定理说明了对于任何具有复本征值 λ 的 2×2 矩阵，例题 7 中的计算都适用. 其证明利用了下列事实：如果 A 中元素是实数，则 $A(\operatorname{Re} x) = \operatorname{Re} Ax$，且 $A(\operatorname{Im} x) = \operatorname{Im} Ax$；如果 x 是对应于复本征值的复本征向量，则 $\operatorname{Re} x$ 和 $\operatorname{Im} x$ 在 $\mathbb{R}^2$ 中线性无关（见习题 25 和 26）. 证明细节略.

【定理 9】 设 A 为 2×2 矩阵，$\lambda = a - bi\,(b\neq0)$ 为复本征值，且 v 为 $\mathbb{C}^2$ 中对应的本征向量. 则 $A = PCP^{-1}$，其中 $P = [\operatorname{Re} v \quad \operatorname{Im} v]$，$C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

例题 7 中的事实在高维空间中也成立. 例如，如果 A 为 3×3 矩阵且有复本征值，则存在 $\mathbb{R}^3$ 中的一个平面，A 在其上的作用为旋转（可能会复合上数乘）. 该平面中的每个向量都旋转为同一平面中的另一点. 称该平面在 A 下**不变**（invariant）.

【例题 8】 矩阵 $A = \begin{bmatrix} 0.8 & -0.6 & 0 \\ 0.6 & 0.8 & 0 \\ 0 & 0 & 1.07 \end{bmatrix}$ 的本征值为 $0.8 \pm 0.6i$ 和 1.07. x_1x_2 平面（第三个坐标为 0）中任一向量 w_0 都被 A 旋转为平面上另一点. 该平面以外的任一向量 x_0 的 x_3 坐标都乘以了 1.07. 点 $w_0 = (2,0,0)$ 和 $x_0 = (2,0,1)$ 在乘以 A 作用下的迭代结果如图 5-15 所示. ■

340

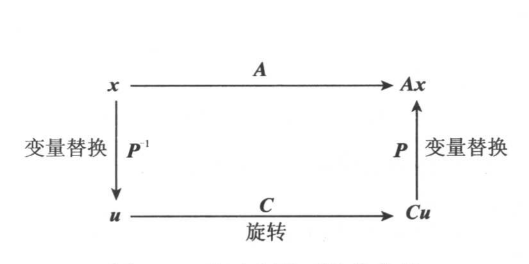


图 5-14　复本征值引发的旋转

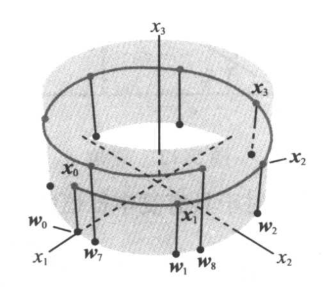


图 5-15　两个点在具有复本征值的 3×3 矩阵作用下的迭代

基础练习

证明：如果 a 和 b 为实数，则 $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ 的本征值为 $a \pm bi$，对应的本征向量为 $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ 和 $\begin{bmatrix} 1 \\ i \end{bmatrix}$.

习题 5.5

令习题 1 ~ 6 中的矩阵作用在 $\mathbf{C}^2$ 上. 试求矩阵的本征值和每个 $\mathbf{C}^2$ 本征空间的一组基.

1. $\begin{bmatrix}1 & -2\\1 & 3\end{bmatrix}$ 2. $\begin{bmatrix}5 & -5\\1 & 1\end{bmatrix}$ 3. $\begin{bmatrix}1 & 5\\-2 & 3\end{bmatrix}$ 4. $\begin{bmatrix}5 & -2\\1 & 3\end{bmatrix}$ 5. $\begin{bmatrix}0 & 1\\-8 & 4\end{bmatrix}$ 6. $\begin{bmatrix}4 & 3\\-3 & 4\end{bmatrix}$

在习题 7 ~ 12 中, 利用例题 6 求 A 的本征值. 在每道题中, 变换 $x \mapsto Ax$ 都是旋转与数乘的复合. 试求旋转的角度 φ, 其中 $-\pi<\varphi\leqslant\pi$, 以及数乘因子 r.

7. $\begin{bmatrix}\sqrt{3} & -1\\1 & \sqrt{3}\end{bmatrix}$ 8. $\begin{bmatrix}\sqrt{3} & 3\\-3 & \sqrt{3}\end{bmatrix}$ 9. $\begin{bmatrix}-\sqrt{3}/2 & 1/2\\-1/2 & -\sqrt{3}/2\end{bmatrix}$ 10. $\begin{bmatrix}-5 & -5\\5 & -5\end{bmatrix}$

11. $\begin{bmatrix}0.1 & 0.1\\-0.1 & 0.1\end{bmatrix}$ 12. $\begin{bmatrix}0 & 0.3\\-0.3 & 0\end{bmatrix}$

在习题 13 ~ 20 中, 试求可逆矩阵 P 和形如 $\begin{bmatrix}a & -b\\b & a\end{bmatrix}$ 的矩阵 C, 使得所给的矩阵具有 $A=PCP^{-1}$ 的形式. 对于习题 13 ~ 16, 可利用习题 1 ~ 4 中的信息.

13. $\begin{bmatrix}1 & -2\\1 & 3\end{bmatrix}$ 14. $\begin{bmatrix}5 & -5\\1 & 1\end{bmatrix}$ 15. $\begin{bmatrix}1 & 5\\-2 & 3\end{bmatrix}$ 16. $\begin{bmatrix}5 & -2\\1 & 3\end{bmatrix}$

17. $\begin{bmatrix}1 & -0.8\\4 & -2.2\end{bmatrix}$ 18. $\begin{bmatrix}1 & -1\\0.4 & 0.6\end{bmatrix}$ 19. $\begin{bmatrix}1.52 & -0.7\\0.56 & 0.4\end{bmatrix}$ 20. $\begin{bmatrix}-1.64 & -2.4\\1.92 & 2.2\end{bmatrix}$

21. 在例题 2 中, 求解 (2) 中第一个方程并用 x_1 表示 x_2, 进而得到 A 的本征向量 $y=\begin{bmatrix}2\\-1+2i\end{bmatrix}$. **证明**: y 是例题 2 中用到的向量 v_2 的 (复) 倍数.

22. 设 A 为复 (或实) $n\times n$ 矩阵, $\mathbf{C}^n$ 中向量 x 为对应于 $\mathbf{C}$ 中本征值 λ 的本征向量. 证明: 对于每个非零复数量 μ, 向量 μx 都为 A 的本征向量.

第7 章将重点研究满足 $A^T=A$ 的矩阵 A. 习题 23 和 24 说明这种矩阵的每个本征值必然为实数.

23. 设 A 为 $n\times n$ 矩阵且 $A^T=A$. x 为 $\mathbf{C}^n$ 中向量, 令 $q=\bar{x}^TAx$. 下面的方程通过验证 $\bar{q}=q$ 说明了 q 为实数. 给出每个等号成立的理由.

$$\bar{q}=\overline{\bar{x}^TAx}=x^T\overline{Ax}=x^TA\bar{x}=(x^TA\bar{x})^T=\bar{x}^TA^Tx=q$$

(a) (b) (c) (d) (e)

24. 设 A 为 $n\times n$ 矩阵且 $A^T=A$. 证明如果对于 $\mathbf{C}^n$ 中某个非零向量 x 有 $Ax=\lambda x$, 则 λ 为实数并且 x 的实部为 A 的本征向量. [提示: 利用习题 23 计算 $\bar{x}^TAx$. 另外, 检验 Ax 的实部和虚部.]

25. 设 A 为 $n\times n$ 实矩阵, x 为 $\mathbf{C}^n$ 中向量. 证明 $\mathrm{Re}(Ax)=A(\mathrm{Re}\,x)$ 且 $\mathrm{Im}(Ax)=A(\mathrm{Im}\,x)$.

26. 设 A 为 2×2 实矩阵, $\lambda=a-bi\,(b\neq 0)$ 为其复本征值, v 为 $\mathbf{C}^n$ 中对应的本征向量.
    a. 证明 $A(\mathrm{Re}\,v)=a\mathrm{Re}\,v+b\mathrm{Im}\,v$ 且 $A(\mathrm{Im}\,v)=-b\,\mathrm{Re}\,v+a\,\mathrm{Im}\,v$. [提示: $v=\mathrm{Re}\,v+i\mathrm{Im}\,v$, 计算 Av.]
    b. 验证如果 P 和 C 如定理 9 中所示, 则 $AP=PC$.

341

[M] 在习题 27 和 28 中, 试对所给的矩阵 A, 求 $A=PCP^{-1}$ 形式的分解, 其中 C 为分块 - 对角矩阵, C 中分块矩阵为形如例题 6 中的 2×2 矩阵. (对于每一对共轭本征值, 利

用 $\mathbf{C}^4$ 中一个本征向量的实部和虚部来构造 $\boldsymbol{P}$ 的两个列向量.)

27. $\begin{bmatrix} 0.7 & 1.1 & 2.0 & 1.7 \\ -2.0 & -4.0 & -8.6 & -7.4 \\ 0 & -0.5 & -1.0 & -1.0 \\ 1.0 & 2.8 & 6.0 & 5.3 \end{bmatrix}$

28. $\begin{bmatrix} -1.4 & -2.0 & -2.0 & -2.0 \\ -1.3 & -0.8 & -0.1 & -0.6 \\ 0.3 & -1.9 & -1.6 & -1.4 \\ 2.0 & 3.3 & 2.3 & 2.6 \end{bmatrix}$

基础练习答案

容易检验一个向量是否为本征向量. 无需考察特征方程. 计算

$$A\boldsymbol{x}=\begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}=\begin{bmatrix} a+b\mathrm{i} \\ b-a\mathrm{i} \end{bmatrix}=(a+b\mathrm{i})\begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}$$

于是$\begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}$为对应于 $\lambda=a+b\mathrm{i}$ 的本征向量. 根据本节的讨论知, $\begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}$一定是对应于 $\bar{\lambda}=a-b\mathrm{i}$ 的本征向量.

5.6 离散动态系统

要理解由差分方程 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 所描述的动态系统的长期行为或演化, 关键在于(掌握)本征值和本征向量. 在1.10节、4.9节和本章的实例介绍中, 我们分别使用了这样的差分方程建立了人口迁移模型、各种不同的马尔可夫链和斑点猫头鹰的种群模型. 向量 $\boldsymbol{x}_k$ 给出了系统随时间(记作 k)变化的信息. 例如, 在斑点猫头鹰的例子中, $\boldsymbol{x}_k$ 列出了在第 k 年, 处于三个不同阶段的猫头鹰的数量.

本节中的应用主要针对生态问题, 因为相对于物理问题或工程问题, 它们更容易说明和解释. 然而, 动态系统在许多科学领域中都会出现. 例如, 在控制系统的标准大学本科课程中会讨论动态系统的若干方面. 这类课程所教授的现代状态-空间设计方法很大程度上依赖于矩阵代数.[1] 控制系统中的稳态响应, 也就是动态系统 $\boldsymbol{x}_{k+1}=A\boldsymbol{x}_k$ 的"长期行为"在工程上的另一种表述.

342

在例题6之前, 都假设 A 可对角化, $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$ 为其线性无关的本征向量, 对应的本征值为 $\lambda_1,\cdots,\lambda_n$. 为了方便, 假设本征向量重新排列使得 $|\lambda_1|\geqslant|\lambda_2|\geqslant\cdots\geqslant|\lambda_n|$. 因为 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$ 为 $\mathbf{R}^n$ 的一组基, 任一初始向量 $\boldsymbol{x}_0$ 可以唯一地表示为

$$\boldsymbol{x}_0=c_1\boldsymbol{v}_1+\cdots+c_n\boldsymbol{v}_n \tag{1}$$

$\boldsymbol{x}_0$ 的这个本征向量分解决定了序列 $\{\boldsymbol{x}_k\}$ 的行为. 下面的计算推广了5.2节例题5所验证的简单情形. 因为 $\boldsymbol{v}_i$ 是本征向量,

$$\boldsymbol{x}_1=A\boldsymbol{x}_0=c_1A\boldsymbol{v}_1+\cdots+c_nA\boldsymbol{v}_n=c_1\lambda_1\boldsymbol{v}_1+\cdots+c_n\lambda_n\boldsymbol{v}_n$$

一般地,

$$\boldsymbol{x}_k=c_1(\lambda_1)^k\boldsymbol{v}_1+\cdots+c_n(\lambda_n)^k\boldsymbol{v}_n \quad (k=0,1,2\cdots) \tag{2}$$

下文的例题将解释当 $k\to\infty$ 时, (2)中可能发生的变化.

5.6.1 捕食者-被捕食者系统

在加利福尼亚的红杉森林深处, 斑点猫头鹰是褐脚森林鼠的主要捕食者, 其多

1. 见 G. F. Franklin, J. D. Powell 和 A. Emami-Naeimi, *Feedback Control of Dynamic Systems*, 4th ed. (Upper Saddle River, NJ: Prentice-Hall, 2001). 这本大学本科教科书对动态模型有很好的介绍(第2章). 状态-空间的设计包含在第6章和第8章.

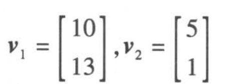
达 80% 的食物都来源于森林鼠. 例题 1 用线性动态系统建立猫头鹰与森林鼠的自然系统模型.（尽管该模型在许多方面都不符合实际,但以此为起点,我们可以进一步研究生态学家所用的一些更为复杂的非线性模型.）

【例题 1】 记猫头鹰和森林鼠在时刻 k 的数量为 $\boldsymbol{x}_k=\begin{bmatrix}O_k\\R_k\end{bmatrix}$，其中 k 是以月份为单位的时间，O_k是研究区域中猫头鹰的数量，R_k是老鼠的数量(单位:千). 假设

$$\begin{aligned}O_{k+1}&=(0.5)O_k+(0.4)R_k\\R_{k+1}&=-p\cdot O_k+(1.1)R_k\end{aligned}\tag{3}$$

其中 p 是一个待定的正参数. 第一个方程中的 $(0.5)O_k$ 说明，如果没有森林鼠做食物，每个月只有一半的猫头鹰可以存活，第二个方程中的 $(1.1)R_k$ 说明，如果没有猫头鹰作为捕食者，老鼠的数量每个月会增加 10%. 如果老鼠充足，猫头鹰的数量将会增加 $(0.4)R_k$，负项 $-p\cdot O_k$ 用以度量猫头鹰的捕食所导致的老鼠的死亡数（事实上,平均每个月被一只猫头鹰吃掉的老鼠的平均数量是 $1000p$）. 当捕食参数 p 是 0.104 时，试确定该系统的演化.

解： 当 $p=0.104$ 时，(3) 的系数矩阵 $\boldsymbol{A}$ 的本征值为 $\lambda_1=1.02$ 和 $\lambda_2=0.58$. 对应的本征向量为



$$\boldsymbol{v}_1=\begin{bmatrix}10\\13\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}5\\1\end{bmatrix}$$

343

初始向量 $\boldsymbol{x}_0$ 可以写为 $\boldsymbol{x}_0=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2$. 于是，对于 $k\geqslant 0$，

$$\boldsymbol{x}_k=c_1(1.02)^k\boldsymbol{v}_1+c_2(0.58)^k\boldsymbol{v}_2=c_1(1.02)^k\begin{bmatrix}10\\13\end{bmatrix}+c_2(0.58)^k\begin{bmatrix}5\\1\end{bmatrix}$$

当 $k\to\infty$ 时，$(0.58)^k$ 迅速的趋于 0. 假定 $c_1>0$，则对于所有足够大的 k，$\boldsymbol{x}_k$ 近似地等于 $c_1(1.02)^k\boldsymbol{v}_1$，写为

$$\boldsymbol{x}_k\approx c_1(1.02)^k\begin{bmatrix}10\\13\end{bmatrix}\tag{4}$$

k 越大(4)的近似程度越高，所以对于大的 k，

$$\boldsymbol{x}_{k+1}\approx c_1(1.02)^{k+1}\begin{bmatrix}10\\13\end{bmatrix}=(1.02)c_1(1.02)^k\begin{bmatrix}10\\13\end{bmatrix}\approx 1.02\boldsymbol{x}_k\tag{5}$$

(5) 中的近似说明，最后 $\boldsymbol{x}_k$ 的每个元素（猫头鹰和老鼠的数量）几乎每个月都近似增加到原来的 1.02 倍，即有 2% 的月增长率. 根据(4)，$\boldsymbol{x}_k$ 约为 (10,13) 的倍数，所以 $\boldsymbol{x}_k$ 中元素的比值约为 10 比 13. 即，每 10 只猫头鹰对应着约 13 000 只老鼠. ■

例题 1 说明了动态系统 $\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k$ 的两个基本事实，其中 $\boldsymbol{A}$ 为 $n\times n$ 矩阵，它的本征值满足：$|\lambda_1|\geqslant 1$，$|\lambda_j|<1$，对于 $j=2,\cdots,n$，并且 $\boldsymbol{v}_1$ 为对应于 λ_1 的本征向量.

如果$\boldsymbol{x}_0$由(1)给出，其中$c_1 \neq 0$，则对于所有充分大的k，

$$\boldsymbol{x}_{k+1} \approx \lambda_1 \boldsymbol{x}_k \tag{6}$$

且

$$\boldsymbol{x}_k \approx c_1(\lambda_1)^k \boldsymbol{v}_1 \tag{7}$$

我们可以选取充分大的k，使(6)和(7)中的近似达到任意精度. 由(6)，最后$\boldsymbol{x}_k$每次增长为原来的λ_1倍，所以λ_1决定了系统的最后增长率. 同样，由(7)，对于大的k值，$\boldsymbol{x}_k$中任何两个元素的比值约等于$\boldsymbol{v}_1$中对应元素的比值. $\lambda_1=1$的情形已在5.2节例题5中说明.

5.6.2　解的图像表示

当$\boldsymbol{A}$为2×2矩阵时，系统演化的几何表示是对代数计算的补充. 可以将方程$\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k$看成映射$\boldsymbol{x}\mapsto\boldsymbol{A}\boldsymbol{x}$重复作用在$\mathbf{R}^2$中的初始点$\boldsymbol{x}_0$时，对$\boldsymbol{x}_0$的变化的描述. $\boldsymbol{x}_0$，$\boldsymbol{x}_1$，…的图像称为动态系统的**轨道**(trajectory).

【例题2】　画出动态系统$\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k$的若干轨道，其中

344

$$\boldsymbol{A}=\begin{bmatrix}0.80 & 0\\ 0 & 0.64\end{bmatrix}$$

解：$\boldsymbol{A}$的本征值为0.8和0.64，对应的本征向量为$\boldsymbol{v}_1=\begin{bmatrix}1\\0\end{bmatrix}$和$\boldsymbol{v}_2=\begin{bmatrix}0\\1\end{bmatrix}$. 如果$\boldsymbol{x}_0=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2$，则

$$\boldsymbol{x}_k=c_1(0.8)^k\begin{bmatrix}1\\0\end{bmatrix}+c_2(0.64)^k\begin{bmatrix}0\\1\end{bmatrix}$$

因为当$k\to\infty$时$(0.8)^k$和$(0.64)^k$都趋于0，所以$\boldsymbol{x}_k$当然也趋于$\boldsymbol{0}$. 但$\boldsymbol{x}_k$趋向$\boldsymbol{0}$的方式十分有趣. 图5-16展示了几条轨道的前几项，这些轨道的起点在以$(\pm3,\pm3)$为角点的方框边界上. 为了便于观察轨道，每个轨道上的点都用一条细线连接起来. ■

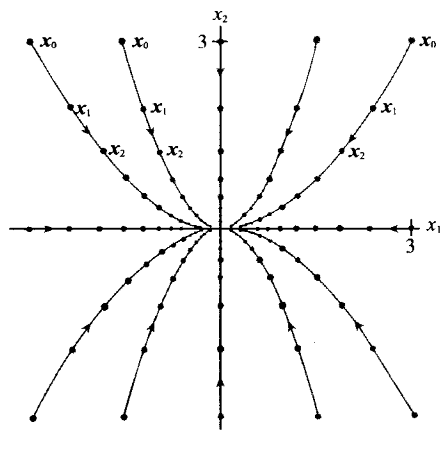


图5-16　原点为吸引子

在例题2中，因为所有的轨道都趋向于**0**，故称原点为动态系统的**吸引子**(attractor)．只要两个本征值的绝对值都小于1，这种情况就会发生．沿过**0**和本征向量$\boldsymbol{v}_2$的这条直线，是吸引最快的方向，其中$\boldsymbol{v}_2$对应的是绝对值较小的本征值．

在下一道例题中，$\boldsymbol{A}$的本征值的绝对值都大于1，且**0**被称为动态系统的**排斥子**(repellor)．除(常量)零以外，$\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k$的所有解都没有边界，且远离原点．[1]

345

【例题3】 画出方程$\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k$的若干典型解，其中

$$\boldsymbol{A}=\begin{bmatrix}1.44 & 0\\ 0 & 1.2\end{bmatrix}$$

解：$\boldsymbol{A}$的本征值为1.44和1.2．如果$\boldsymbol{x}_0=\begin{bmatrix}c_1\\ c_2\end{bmatrix}$，则

$$\boldsymbol{x}_k=c_1(1.44)^k\begin{bmatrix}1\\0\end{bmatrix}+c_2(1.2)^k\begin{bmatrix}0\\1\end{bmatrix}$$

两项的大小都增长，但第一项增长得快一些．所以沿着过**0**和绝对值较大本征值所对应的本征向量的这条直线，是排斥最快的方向．图5-17展示了几个轨道，它们的起点都非常接近**0**． ■

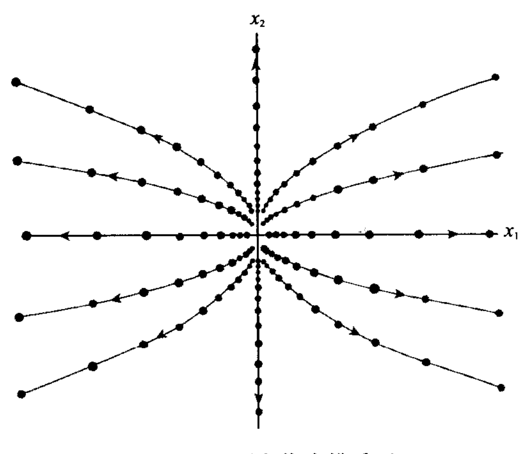


图5-17 原点作为排斥子

在下一道例题中，因为原点在某些方向吸引解，但在其他方向又排斥它们，故称**0**为**鞍点**(saddle point)．只要一个本征值的绝对值大于1且另外一个绝对值小于1，这种情况就会发生．吸引最快的方向由较小本征值对应的本征向量确定．排斥最快的方向由较大本征值对应的本征向量确定．

1. 在线性动态系统中，可能的吸引子或者排斥子只有原点，但在一个更一般的动态系统中，当映射$\boldsymbol{x}_{k+1}\mapsto\boldsymbol{x}_k$不是线性时，存在多个吸引子和排斥子．在这样的系统中，吸引子和排斥子是通过一个特殊矩阵(有变量为其元素)的本征值定义的，这个矩阵叫做系统的雅可比矩阵．

【例题 4】　画出方程 $\boldsymbol{y}_{k+1}=\boldsymbol{D}\boldsymbol{y}_k$ 的若干典型解，其中

$$\boldsymbol{D}=\begin{bmatrix}2.0 & 0\\ 0 & 0.5\end{bmatrix}$$

（这里我们用 $\boldsymbol{D}$ 和 $\boldsymbol{y}$ 代替 $\boldsymbol{A}$ 和 $\boldsymbol{x}$ 是因为这个例题后面还会用到）. 证明：如果解 346 $\{\boldsymbol{y}_k\}$的起点不在 x_2轴上，则该解无边界.

解：$\boldsymbol{D}$ 的本征值为 2 和 0.5. 如果 $\boldsymbol{y}_0=\begin{bmatrix}c_1\\ c_2\end{bmatrix}$，则

$$\boldsymbol{y}_k=c_1 2^k\begin{bmatrix}1\\ 0\end{bmatrix}+c_2(0.5)^k\begin{bmatrix}0\\ 1\end{bmatrix} \tag{8}$$

如果 $\boldsymbol{y}_0$在 x_2轴上，则 $c_1=0$ 且当 $k\to\infty$ 时 $\boldsymbol{y}_k\to\boldsymbol{0}$. 但如果 $\boldsymbol{y}_0$不在 x_2轴上，则 $\boldsymbol{y}_k$的和式的第一项可以变得任意大，所以$\{\boldsymbol{y}_k\}$无边界. 图 5-18 展示了 10 个轨道，它们的起点在 x_2轴上或在其附近. ■

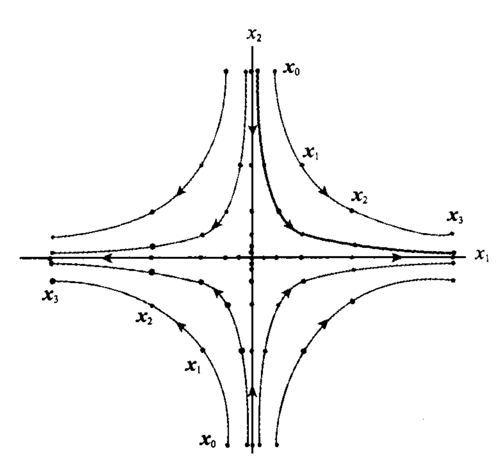


图 5-18　原点为鞍点

5.6.3　变量替换

前面的三个例题都是关于对角矩阵的. 为了处理非对角情形，我们暂时回到 $n\times n$ 的情形，其中 $\boldsymbol{A}$ 的本征向量构成 $\mathbb{R}^n$的一组基$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$. 令 $\boldsymbol{P}=[\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n]$，$\boldsymbol{D}$ 为对角矩阵，其对角线元素为 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$ 对应的本征值. 已知序列$\{\boldsymbol{x}_k\}$满足 $\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k$，定义一个新的序列为

$$\boldsymbol{y}_k=\boldsymbol{P}^{-1}\boldsymbol{x}_k,\text{或等价地},\boldsymbol{x}_k=\boldsymbol{P}\boldsymbol{y}_k$$

将这些等式代入方程 $\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k$ 并且根据 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$的事实，得到

$$\boldsymbol{P}\boldsymbol{y}_{k+1}=\boldsymbol{A}\boldsymbol{P}\boldsymbol{y}_k=(\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1})\boldsymbol{P}\boldsymbol{y}_k=\boldsymbol{P}\boldsymbol{D}\boldsymbol{y}_k$$

两边同时左乘 $\boldsymbol{P}^{-1}$，得到

347

$$\boldsymbol{y}_{k+1}=\boldsymbol{D}\boldsymbol{y}_k$$

如果我们将 $\boldsymbol{y}_k$写为 $\boldsymbol{y}(k)$，并且记 $\boldsymbol{y}(k)$中的元素为 $y_1(k),\cdots,y_n(k)$，则

$$\begin{bmatrix} y_1(k+1) \\ y_2(k+1) \\ \vdots \\ y_n(k+1) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_n(k) \end{bmatrix}$$

从$\boldsymbol{x}_k$到$\boldsymbol{y}_k$的变量替换实现了对差分方程组的解耦. 例如，$y_1(k)$的变化不受$y_2(k),\cdots,y_n(k)$的影响，因为对每个k都有，$y_1(k+1)=\lambda_1 y_1(k)$.

方程$\boldsymbol{x}_k=\boldsymbol{P}\boldsymbol{y}_k$说明$\boldsymbol{y}_k$是$\boldsymbol{x}_k$关于本征向量基$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$的坐标向量. 我们可以通过在新的本征向量坐标系中的计算来解耦系统$\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k$. 当$n=2$时，这相当于使用以两个本征向量方向为坐标轴的坐标纸.

【例题 5】 证明原点是$\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k$的解的鞍点，其中

$$\boldsymbol{A}=\begin{bmatrix} 1.25 & -0.75 \\ -0.75 & 1.25 \end{bmatrix}$$

并且求出吸引最快和排斥最快的方向.

解：通过标准的演算可以得到：$\boldsymbol{A}$的本征值为 2 和 0.5，对应的本征向量分别为$\boldsymbol{v}_1=\begin{bmatrix} 1 \\ -1 \end{bmatrix}$和$\boldsymbol{v}_2=\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. 因为$|2|>1$且$|0.5|<1$，所以原点为动态系统的鞍点. 如果$\boldsymbol{x}_0=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2$，则

$$\boldsymbol{x}_k=c_1 2^k \boldsymbol{v}_1+c_2(0.5)^k \boldsymbol{v}_2 \tag{9}$$

这个方程看上去与例题 4 中的(8)很相似，只不过这里用$\boldsymbol{v}_1,\boldsymbol{v}_2$代替了标准基.

在坐标纸上，画出过$\boldsymbol{0}$和本征值$\boldsymbol{v}_1$和$\boldsymbol{v}_2$的坐标轴. 见图 5-19. 沿着这些坐标轴的运动相当于图 5-18 中沿着标准坐标轴的运动. 在 5-19 中，排斥最快的方向是过$\boldsymbol{0}$和本征向量$\boldsymbol{v}_1$的直线方向，其中$\boldsymbol{v}_1$对应的本征值绝对值大于 1. 如果$\boldsymbol{x}_0$在这条直线上，则(9)中的c_2为零并且$\boldsymbol{x}_k$迅速远离$\boldsymbol{0}$. 吸引最快的方向由本征向量$\boldsymbol{v}_2$确定，其对应的本征值的绝对值小于 1.

图 5-19 展示了很多轨道. 从本征向量坐标轴的角度观察该图时，它“看上去”和图 5-18 完全一样. ■

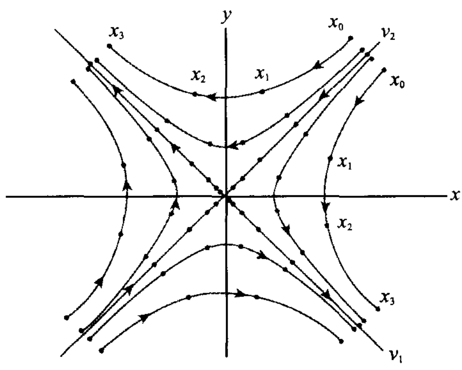


图 5-19 原点为鞍点

5.6.4　复本征值

当2×2矩阵$\boldsymbol{A}$有复本征值时，$\boldsymbol{A}$不可对角化(当$\boldsymbol{A}$作用在$\mathbf{R}^n$上时)，但动态系统$\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k$却容易描述. 5.5节中例题3说明了本征值绝对值为1的情形. 点$\boldsymbol{x}_0$的迭代形成环绕原点的一条椭圆形轨道.

348

如果$\boldsymbol{A}$有两个绝对值大于1的复本征值，则$\mathbf{0}$为排斥子，且$\boldsymbol{x}_0$的迭代将会环绕原点向外盘旋. 如果复本征的绝对值小于1，则原点为吸引子，且$\boldsymbol{x}_0$的迭代将会环绕原点向内盘旋，如下面的例题所示.

【例题6】　可以验证矩阵

$$\boldsymbol{A}=\begin{bmatrix}0.8 & 0.5\\ -0.1 & 1.0\end{bmatrix}$$

的本征值为$0.9\pm0.2\text{i}$,对应的本征向量为$\begin{bmatrix}1\mp2\text{i}\\ 1\end{bmatrix}$. 图5-20展示了系统$\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k$的三条轨道,起始向量分别为$\begin{bmatrix}0\\ 2.5\end{bmatrix}$,$\begin{bmatrix}3\\ 0\end{bmatrix}$和$\begin{bmatrix}0\\ -2.5\end{bmatrix}$. ■

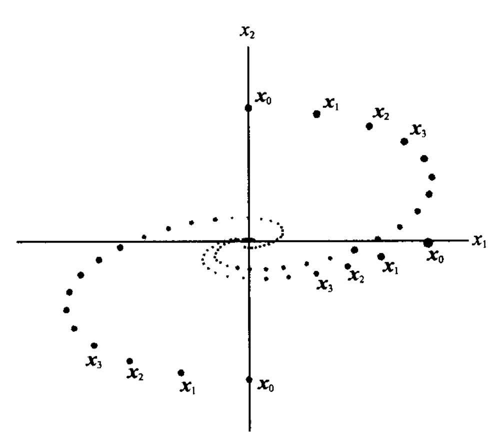


图5-20　复本征向量对应的旋转

5.6.5　斑点猫头鹰的生存

在本章的实例介绍中,通过动态系统$\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k$建立了加利福尼亚柳河湾地区斑点猫头鹰的种群模型. $\boldsymbol{x}_k=(j_k,s_k,a_k)$分别列出了处于幼鸟期、成长期和成熟期的雌猫头鹰的数量，$\boldsymbol{A}$为状态－矩阵

349

$$\boldsymbol{A}=\begin{bmatrix}0 & 0 & 0.33\\ 0.18 & 0 & 0\\ 0 & 0.71 & 0.94\end{bmatrix}\tag{10}$$

用MATLAB可以求出$\boldsymbol{A}$的本征值约为$\lambda_1=0.98$,$\lambda_2=-0.02+0.21\text{i}$和$\lambda_3=-0.02-0.21\text{i}$. 因为$|\lambda_2|^2=|\lambda_3|^2=(-0.02)^2+(0.21)^2=0.0445$，所以三个本征值的绝对值都小于1.

此时，令 A 作用在复向量空间 C^3 上. 因为 A 有三个不同的本征值，所以三个对应的本征向量线性无关且构成 C^3 的一组基. 记本征向量为 v_1, v_2 和 v_3. 因此 $x_{k+1}=Ax_k$ 的通解形如(使用 C^3 中向量)

$$x_k=c_1(\lambda_1)^k v_1+c_2(\lambda_2)^k v_2+c_3(\lambda_3)^k v_3 \tag{11}$$

如果 x_0 为实起始向量，则由于 A 为实矩阵，所以 $x_1=Ax_0$ 为实向量. 类似地，方程 $x_{k+1}=Ax_k$ 说明(11)左边的每个 x_k 都是实向量，尽管它是复向量的和. 因为所有的本征值的绝对值都小于1，所以(11)右边的每一项趋于零向量. 因此实序列$\{x_k\}$也趋于零向量. 非常遗憾，这个模型预示着斑点猫头鹰最终将会灭绝.

斑点猫头鹰还有生存的可能吗？回忆实例介绍，(10)中矩阵 A 的元素 18% 基于下列事实：尽管有 60% 的幼鸟可以离开旧巢并寻找新领地，但其中只有 30% 能在寻找的过程中存活并找到自己的新领地. 寻找过程中的存活率主要受森林中皆伐区的数量影响，因为它使得寻找新领地更加困难和危险.

有些猫头鹰种群生活在没有或少有皆伐区的区域中. 在那里，幼鸟得以存活并找到新领地的比率更高一些. 当然，斑点猫头鹰的问题比我们描述的更加复杂，但最后一道例题将为这件事画上完满的句号. 350

【例题7】 假设猫头鹰幼鸟在寻找过程中的存活率为50%，则(10)中状态-矩阵 A 的(2,1)-元是 0.3 而不是 0.18. 对于该斑点猫头鹰的种群数量，这个状态-矩阵模型预示着什么？

解： 现在 A 的本征值约为 $\lambda_1=1.01$, $\lambda_2=-0.03+0.26i$ 和 $\lambda_3=-0.03-0.26i$. λ_1 对应的本征向量约为 $v_1=(10,3,31)$. 设 λ_2 和 λ_3 对应的(复)本征向量为 v_2和 v_3. 此时，方程(11)变为

$$x_k=c_1(1.01)^k v_1+c_2(-0.03+0.26i)^k v_2+c_3(-0.03-0.26i)^k v_3$$

当 $k\to\infty$ 时，后两个向量趋于零. 所以 x_k 变得越来越接近(实)向量 $c_1(1.01)^k v_1$. 这恰好是例题1下面(6)和(7)中的两个近似在本例的应用. 此外，还可以证明当初始向量 x_0的元素非负时，常数 c_1为正. 因此猫头鹰的数量将缓慢增加，其长期增长率为 1.01. 按照其生长阶段，本征向量 v_1描述了猫头鹰的最终分布：对于每 31 只处于幼鸟期的猫头鹰，大约有 10 只处于成熟期和 3 只处于成长期的猫头鹰与之对应. ■

更多阅读

Franklin, G. F., I. D. Powell, and M. L. Workman. *Digital Control of Dynamic Systems*, 3rd ed. Reading, MA: Addison-Wesley, 1998.

Sandefur, James T. *Discrete Dynamical Systems - Theory and Applications*. Oxford: Oxford University Press, 1990.

Tuchinsky, Philip. *Management of a Buffalo Herd*, UMAP Module 207. Lexington, MA: COMAP, 1980.

基础练习

1. 下面的矩阵本征值为 1, 2/3 和 1/3，对应的本征向量分别为 v_1, v_2 和 v_3：

$$A=\frac{1}{9}\begin{bmatrix}7&-2&0\\-2&6&2\\0&2&5\end{bmatrix}, v_1=\begin{bmatrix}-2\\2\\1\end{bmatrix}, v_2=\begin{bmatrix}2\\1\\2\end{bmatrix}, v_3=\begin{bmatrix}1\\2\\-2\end{bmatrix}$$

若 $x_0=\begin{bmatrix}1\\11\\-2\end{bmatrix}$，试求方程 $x_{k+1}=Ax_k$ 的通解.

351 2. 当 $k\to\infty$ 时，考察基础练习 1 中序列 $\{x_k\}$ 的变化.

习题　5.6

1. 设 A 为 2×2 矩阵，本征值为 3 和 1/3 且对应的本征向量为 $v_1=\begin{bmatrix}1\\1\end{bmatrix}$ 和 $v_2=\begin{bmatrix}-1\\1\end{bmatrix}$. 设 $\{x_k\}$ 为差分方程 $x_{k+1}=Ax_k$ 的解，$x_0=\begin{bmatrix}9\\1\end{bmatrix}$.
   a. 计算 $x_1=Ax_0$.（提示：无需知道 A 本身.）
   b. 试求 x_k 的表达式，其中包含 k 及本征向量 v_1 和 v_2.
2. 设矩阵 A 的本征值为 3, 4/5, 3/5，对应的本征向量为 $\begin{bmatrix}1\\0\\-3\end{bmatrix}$, $\begin{bmatrix}2\\1\\-5\end{bmatrix}$, $\begin{bmatrix}-3\\-3\\7\end{bmatrix}$. 令 $x_0=\begin{bmatrix}-2\\-5\\3\end{bmatrix}$. 对于指定的 x_0，试求方程 $x_{k+1}=Ax_k$ 的解，并描述当 $k\to\infty$ 时其变化.

在习题 3 ~ 6 中，假设对任一初始向量都存在本征向量分解，使得本节方程(1)中的系数 c_1 为正.[1]

3. 在例题 1 里，当(3)中的捕食参数 p 是 0.2 时，试确定该动态系统的演化（给出 x_k 的计算公式）. 猫头鹰的数量是上升还是下降？森林鼠的数量又如何呢？
4. 当例题 1 中的捕食参数 p 是 0.125 时，试确定该动态系统的演化（给出 x_k 的计算公式）. 猫头鹰和森林鼠的数量随着时间如何变化？该系统趋向一种被称为不稳定平衡的状态. 如果该系统的某个方面（比如出生率或捕食率）有轻微的变动，系统会如何变化？
5. 在美国黄杉森林中，斑点猫头鹰主要以鼯鼠为食. 假设这两个种群的捕食者 - 被捕食者矩阵为 $A=\begin{bmatrix}0.4 & 0.3\\-p & 1.2\end{bmatrix}$. 证明：如果捕食参数 p 为 0.325，则两个种群都会增长. 估计这个长期增长率及猫头鹰与鼯鼠的最终比值.
6. 证明：如果习题 5 中的捕食率为 0.5，则猫头鹰和鼯鼠最终都将灭绝. 试求一个 p 值，使得猫头鹰和鼯鼠的数量都趋于稳定. 此时，对应的种群数量是多少？
7. 设 A 具有习题 1 中描述的性质.
   a. 原点是动态系统 $x_{k+1}=Ax_k$ 的吸引子，排斥子，或者鞍点吗？
   b. 试求该系统吸引最快和/或排斥最快的方向.
   c. 给出该系统的图像表示，指出吸引最快或排斥最快的方向. 绘出一些典型轨道的草图（不计算特殊点）.

---

1. 例题 1 中模型的限制之一是，种群模型总是存在初始向量 x_0，其元素为正且使得系数 c_1 为负. (7)中的近似仍有效，但 x_k 中的元素最终变为负的.

8. 如果 A 具有习题2 中描述的性质，对于动态系统 $x_{k+1}=Ax_k$，试确定原点的属性(吸引子,排斥子,鞍点). 试求吸引或排斥最快的方向.

在习题9~14 中，将动态系统 $x_{k+1}=Ax_k$ 的原点以吸引子、排斥子或鞍点分类. 试求吸引和/或排斥最快的方向.

9. $A=\begin{bmatrix}1.7 & -0.3\\ -1.2 & 0.8\end{bmatrix}$ 10. $A=\begin{bmatrix}0.3 & 0.4\\ -0.3 & 1.1\end{bmatrix}$ 11. $A=\begin{bmatrix}0.4 & 0.5\\ -0.4 & 1.3\end{bmatrix}$

12. $A=\begin{bmatrix}0.5 & 0.6\\ -0.3 & 1.4\end{bmatrix}$ 13. $A=\begin{bmatrix}0.8 & 0.3\\ -0.4 & 1.5\end{bmatrix}$ 14. $A=\begin{bmatrix}1.7 & 0.6\\ -0.4 & 0.7\end{bmatrix}$

15. 设 $A=\begin{bmatrix}0.4 & 0 & 0.2\\ 0.3 & 0.8 & 0.3\\ 0.3 & 0.2 & 0.5\end{bmatrix}$. 向量 $v_1=\begin{bmatrix}0.1\\ 0.6\\ 0.3\end{bmatrix}$为 A 的一个本征向量，且0.5 和0.2 为其两个本征值. 构造动态系统 $x_{k+1}=Ax_k$ 的解，满足 $x_0=(0,0.3,0.7)$. 当 $k\to\infty$ 时，x_k如何变化?

16. [M] 求动态系统 $x_{k+1}=Ax_k$ 的通解，其中 A 为4.9 节例题16 中 Hertz Rent A Car 模型的随机矩阵.

17. 构造一种动物的状态-矩阵模型，这种动物有两个生命阶段：幼体(直到1 岁)和成体. 假设成体的雌性动物平均每年生育 1.6 只雌性幼崽. 每年，有 30% 的幼崽存活并成长为成体，有 80% 的成体能够存活下来. 对于 $k\geqslant 0$，令 $x_k=(j_k,a_k)$，其中 x_k 中元素为第 k 年幼体和成体的数量. 352
   a. 构造状态-矩阵 A 使得对于 $k\geqslant 0$，$x_{k+1}=Ax_k$.
   b. 证明该物种的数量在增长. 求最终种群数量的增长率，以及幼体与成体的数量之比.
   c. [M]假设最初有 15 只幼体动物和 10 只成体动物. 画出四个图来显示在八年内种群的变化：(a)幼体的数量，(b)成体的数量，(c)总的数量，(d)幼体与成体（每年）的数量之比. (d)中的比值何时稳定? 写出用来绘制(c)和(d)中图像的程序或命令.

18. 一群美国水牛(野牛)可以像斑点猫头鹰那样通过状态矩阵来建模. 雌性可以分为产犊(直到1 岁)、一年崽(1 到2 岁)和成体. 假设每 100 头雌成体平均每年生育 42 头雌产犊(只有成体可以生育后代). 每年大约有 60% 的产犊存活，75% 的一年崽存活和 95% 的成体存活. 对于 $k\geqslant 0$，令 $x_k=(c_k,y_k,a_k)$，其中 x_k 的元素为第 k 年各个生长阶段中雌性的数量.
   a. 构造水牛群的阶段-矩阵模型 A，使得对于 $k\geqslant 0$，$x_{k+1}=Ax_k$.
   b. [M]证明水牛群正在增长，试确定最终的增长率，以及每 100 头成体对应的产犊和一年崽的数量.

基础练习答案

1. 第一步将 x_0 写成 v_1，v_2，v_3 的线性组合. 将$[v_1\ \ v_2\ \ v_3\ \ x_0]$行化简得到系数为 $c_1=2$，$c_2=1$，$c_3=3$，所以

$$x_0=2v_1+1v_2+3v_3$$

又因为本征值为 1,2/3 和 1/3，所以通解为

$$x_k=2\times 1^k v_1+1\times\left(\frac{2}{3}\right)^k v_2+3\times\left(\frac{1}{3}\right)^k v_3=2\begin{bmatrix}-2\\ 2\\ 1\end{bmatrix}+\left(\frac{2}{3}\right)^k\begin{bmatrix}2\\ 1\\ 2\end{bmatrix}+3\times\left(\frac{1}{3}\right)^k\begin{bmatrix}1\\ 2\\ -2\end{bmatrix}\quad(12)$$

2. 当 $k\to\infty$ 时,(12)中的第二项和第三项趋于零向量，并且

$$x_k=2v_1+\left(\frac{2}{3}\right)^k v_2+3\left(\frac{1}{3}\right)^k v_3\to 2v_1=\begin{bmatrix}-4\\ 4\\ 2\end{bmatrix}$$

5.7 微分方程中的应用

本节介绍的微分方程是5.6节中差分方程在连续情形下的类推. 在许多应用问题中，一些量会随时间连续地变化，并且与下列微分方程组有关：

353

$$\begin{aligned} x'_1 &= a_{11}x_1 + \cdots + a_{1n}x_n \\ x'_2 &= a_{21}x_1 + \cdots + a_{2n}x_n \\ &\vdots \\ x'_n &= a_{n1}x_1 + \cdots + a_{nn}x_n \end{aligned}$$

其中 $x_1,\cdots,x_n$ 为 t 的可微函数，$x'_1,\cdots,x'_n$ 为其导数，a_{ij} 为常量. 这个系统重要的特性是它是线性的. 为了说明这一点，将该方程组写成矩阵微分方程的形式

$$\boldsymbol{x}' = A\boldsymbol{x} \tag{1}$$

其中

$$\boldsymbol{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad \boldsymbol{x}'(t) = \begin{bmatrix} x'_1(t) \\ \vdots \\ x'_n(t) \end{bmatrix}, \quad 且 A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

方程(1)的一个**解**(solution)是指对于某个实数区间内所有的 t，如 $t \geqslant 0$，都满足(1)的一个向量值函数.

方程(1)是线性的，因为函数求导和矩阵与向量乘法都是线性变换. 从而，如果 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是 $\boldsymbol{x}' = A\boldsymbol{x}$ 的解，则 $c\boldsymbol{u} + d\boldsymbol{v}$ 也是解，因为

$$(c\boldsymbol{u} + d\boldsymbol{v})' = c\boldsymbol{u}' + d\boldsymbol{v}' = cA\boldsymbol{u} + dA\boldsymbol{v} = A\ (c\boldsymbol{u} + d\boldsymbol{v})$$

工程师们将这个性质称为解的*叠加*. 同样，恒等于零的函数是(1)的(平凡)解. 在第4章的概念中，(1)所有解的集合是 $\mathbb{R}^n$中所有连续实值函数构成的集合的子空间.

关于微分方程的标准教材会指出(1)的**基础解系**(fundamental set of solutions)总是存在的. 如果 A 为 $n \times n$ 矩阵，则基础解系中有 n 个线性无关的函数，并且(1)的每个解都是这 n 个函数的唯一线性组合. 也就是说，基础解系是(1)的解集的基，并且解集是 n 维函数向量空间. 如果指定向量 $\boldsymbol{x}_0$，则**初值问题**(initial value problem)就是构造(唯一的)函数 $\boldsymbol{x}$ 使得 $\boldsymbol{x}' = A\boldsymbol{x}$ 且 $\boldsymbol{x}(0) = \boldsymbol{x}_0$.

当 A 为对角矩阵，(1)的解可以由初等微积分求出. 例如，考虑

$$\begin{bmatrix} x'_1(t) \\ x'_2(t) \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{2}$$

即

$$\begin{aligned} x'_1(t) &= \ \ 3x_1(t) \\ x'_2(t) &= -5x_2(t) \end{aligned} \tag{3}$$

方程组(2)称为解耦系统，因为每个函数的导数仅依赖于函数本身，而不依赖于$x_1(t)$和$x_2(t)$的组合或“耦合”. 根据微积分的知识，(3)的解为 $x_1(t) = c_1\mathrm{e}^{3t}$和$x_2(t) = c_2\mathrm{e}^{-5t}$，对于任何常数 c_1和 c_2. (2)的每个解都可写成如下形式

354

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} c_1\mathrm{e}^{3t} \\ c_2\mathrm{e}^{-5t} \end{bmatrix} = c_1\begin{bmatrix} 1 \\ 0 \end{bmatrix}\mathrm{e}^{3t} + c_2\begin{bmatrix} 0 \\ 1 \end{bmatrix}\mathrm{e}^{-5t}$$

这个例子表明对于一般的方程 $\boldsymbol{x}'=A\boldsymbol{x}$，其解可能为如下形式函数的线性组合

$$\boldsymbol{x}(t)=\boldsymbol{v}e^{\lambda t} \tag{4}$$

对于某个数量 λ 和某个固定非零向量 $\boldsymbol{v}$. [如果 $\boldsymbol{v}=\boldsymbol{0}$，函数 $\boldsymbol{x}(t)$ 恒等于零，从而满足 $\boldsymbol{x}'=A\boldsymbol{x}$.]注意到

$$\boldsymbol{x}'(t)=\lambda\boldsymbol{v}e^{\lambda t} \quad \text{由微积分，因为 } \boldsymbol{v} \text{ 为常向量}$$

$$A\boldsymbol{x}(t)=A\boldsymbol{v}e^{\lambda t} \quad \text{(4)两边同时乘以 } A$$

由于 $e^{\lambda t}$ 不等于零，$\boldsymbol{x}'(t)$ 等于 $A\boldsymbol{x}(t)$ 当且仅当 $\lambda\boldsymbol{v}=A\boldsymbol{v}$，即当且仅当 λ 为 A 的本征值，且 $\boldsymbol{v}$ 为对应的本征向量. 因此每一对本征值-本征向量给出了 $\boldsymbol{x}'=A\boldsymbol{x}$ 的一个解(4). 这样的解称为微分方程的**特征函数**. 特征函数是求解微分方程组的关键.

【例题 1】 图 5-21 中的电路可以由下面的微分方程刻画

$$\begin{bmatrix} v'_1(t) \\ v'_2(t) \end{bmatrix}=\begin{bmatrix} -(1/R_1+1/R_2)/C_1 & 1/(R_2C_1) \\ 1/(R_2C_2) & -1/(R_2C_2) \end{bmatrix}\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}$$

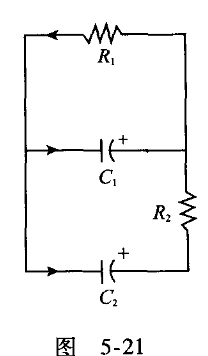


图 5-21

其中 $v_1(t)$ 和 $v_2(t)$ 分别为两个电容器在时刻 t 的电压. 设电阻 R_1 等于 1 欧姆，R_2 等于 2 欧姆，电容 C_1 等于 1 法拉，C_2 等于 0.5 法拉，并且设电容器 C_1 和 C_2 的初始电压分别是 5 伏特和 4 伏特. 试求 $v_1(t)$ 和 $v_2(t)$ 的公式，并描述电压是如何随着时间变化的.

解： 根据已知数据，令 $A=\begin{bmatrix} -1.5 & 0.5 \\ 1 & -1 \end{bmatrix}$，$\boldsymbol{x}=\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$，$\boldsymbol{x}_0=\begin{bmatrix} 5 \\ 4 \end{bmatrix}$.

向量 $\boldsymbol{x}_0$ 给出了 $\boldsymbol{x}$ 的初始值. 由 A，得到本征值 $\lambda_1=-5$ 和 $\lambda_2=-2$，对应的本征向量为

$$\boldsymbol{v}_1=\begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{和 } \boldsymbol{v}_2=\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

特征函数 $x_1(t)=\boldsymbol{v}_1e^{\lambda_1 t}$，$x_2(t)=\boldsymbol{v}_2e^{\lambda_2 t}$ 以及 $\boldsymbol{x}_1$ 与 $\boldsymbol{x}_2$ 的任何线性组合都满足 $\boldsymbol{x}'=A\boldsymbol{x}$. 令

$$\boldsymbol{x}(t)=c_1\boldsymbol{v}_1e^{\lambda_1 t}+c_2\boldsymbol{v}_2e^{\lambda_2 t}=c_1\begin{bmatrix} 1 \\ 2 \end{bmatrix}e^{-0.5t}+c_2\begin{bmatrix} -1 \\ 1 \end{bmatrix}e^{-2t}$$

并且注意到 $\boldsymbol{x}(0)=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2$. 显然，$\boldsymbol{v}_1, \boldsymbol{v}_2$ 线性无关因此可以生成 $\mathbf{R}^2$，并且可以求出 c_1 和 c_2 使得 $\boldsymbol{x}(0)$ 等于 $\boldsymbol{x}_0$. 事实上，由方程

355

$$c_1\begin{bmatrix} 1 \\ 2 \end{bmatrix}+c_2\begin{bmatrix} -1 \\ 1 \end{bmatrix}=\begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

$$\uparrow \qquad\qquad \uparrow \qquad\qquad \uparrow$$

$$\boldsymbol{v}_1 \qquad\qquad \boldsymbol{v}_2 \qquad\qquad \boldsymbol{x}_0$$

易得 $c_1=3$ 且 $c_2=-2$. 所以微分方程 $\boldsymbol{x}'=A\boldsymbol{x}$ 的解为

$$\boldsymbol{x}(t)=3\begin{bmatrix} 1 \\ 2 \end{bmatrix}e^{-0.5t}-2\begin{bmatrix} -1 \\ 1 \end{bmatrix}e^{-2t}$$

或

$$\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}=\begin{bmatrix} 3e^{-0.5t}+2e^{-2t} \\ 6e^{-0.5t}-2e^{-2t} \end{bmatrix}$$

图 5-22 是 $t\geqslant 0$ 时，$\boldsymbol{x}(t)$ 的图像或者是轨道，和轨道一起被绘出的还有其他一些初始点. 两个特征函数 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 的轨道在 $\boldsymbol{A}$ 的本征空间中.

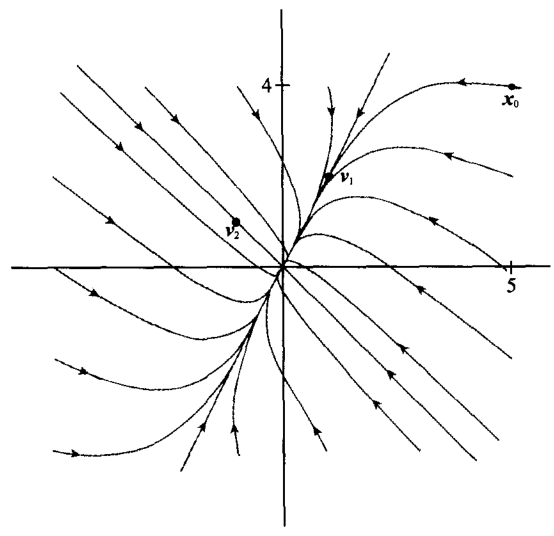


图 5-22　原点为吸引子

当 $t\to\infty$ 时，函数 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 都衰减至零，但因为 $\boldsymbol{x}_2$ 的负指数较小，所以它的值衰退的速度较快. 对应的本征向量 $\boldsymbol{v}_2$ 中的元素说明，如果初始电压绝对值相等但符号相反，则电容器所带的电压将迅速的衰减至零. ■

在图 5-22 中，因为全体轨道都趋于原点，所以原点称为动态系统的**吸引子**(attractor)，或**汇点**(sink). 吸引最快的方向是沿着特征函数 $\boldsymbol{x}_2$ 的轨道(沿着通过 $\boldsymbol{0}$ 和 $\boldsymbol{v}_2$ 的直线)，其中特征函数 $\boldsymbol{x}_2$ 对应于较小的负本征值 $\lambda=-2$. 起点在直线以外的轨道，其渐进线是过 $\boldsymbol{0}$ 和 $\boldsymbol{v}_1$ 的直线，因为它们在 $\boldsymbol{v}_2$ 方向上的分量迅速的衰减.

356

如果例题 1 中的本征值为正而不是负，则对应的轨道形状是相似的，但是轨道将远离原点. 在这种情形下，原点称为动态系统的**排斥子**(repellor)，或**源点**(source)，排斥最快的方向是包含特征函数的直线，这个特征函数对应于较大的正本征值.

【例题 2】　假设一个质点在平面力场中运动，其位置向量 $\boldsymbol{x}$ 满足 $\boldsymbol{x}'=\boldsymbol{A}\boldsymbol{x}$ 和 $\boldsymbol{x}(0)=\boldsymbol{x}_0$，其中

$$\boldsymbol{A}=\begin{bmatrix}4 & -5\\ -2 & 1\end{bmatrix},\quad \boldsymbol{x}_0=\begin{bmatrix}2.9\\ 2.6\end{bmatrix}$$

对 $t\geqslant 0$ 求解这个初值问题，并且绘出质点轨道的草图.

解： $\boldsymbol{A}$ 的本征值为 $\lambda_1=6$ 和 $\lambda_2=-1$，对应的本征向量为 $\boldsymbol{v}_1=(-5,2)$ 和 $\boldsymbol{v}_2=(1,1)$. 对于任意常量 c_1 和 c_2，函数

$$\boldsymbol{x}(t)=c_1\boldsymbol{v}_1\mathrm{e}^{\lambda_1 t}+c_2\boldsymbol{v}_2\mathrm{e}^{\lambda_2 t}=c_1\begin{bmatrix}-5\\ 2\end{bmatrix}\mathrm{e}^{6t}+c_2\begin{bmatrix}1\\ 1\end{bmatrix}\mathrm{e}^{-t}$$

是 $\boldsymbol{x}'=\boldsymbol{A}\boldsymbol{x}$ 的解. 要求 c_1 和 c_2 使得 $\boldsymbol{x}(0)=\boldsymbol{x}_0$，即

$$c_1\begin{bmatrix}-5\\2\end{bmatrix}+c_2\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}2.9\\2.6\end{bmatrix}\text{或}\begin{bmatrix}-5&1\\2&1\end{bmatrix}\begin{bmatrix}c_1\\c_2\end{bmatrix}=\begin{bmatrix}2.9\\2.6\end{bmatrix}$$

求解得 $c_1=-3/70$ 和 $c_2=188/70$，故所求函数为

$$\boldsymbol{x}(t)=\frac{-3}{70}\begin{bmatrix}-5\\2\end{bmatrix}e^{6t}+\frac{188}{70}\begin{bmatrix}1\\1\end{bmatrix}e^{-t}$$

$\boldsymbol{x}$ 和其他解的轨道如图 5-23 所示. ■

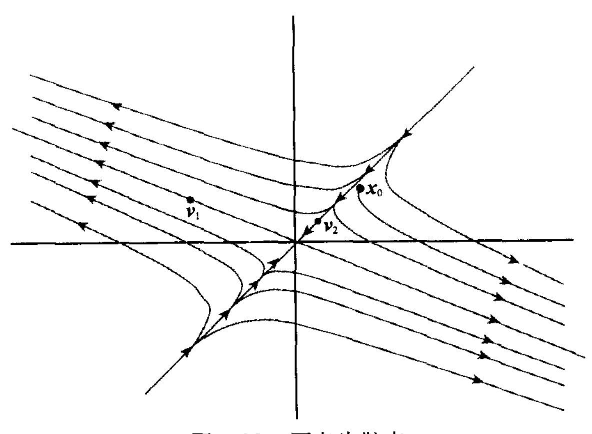


图 5-23 原点为鞍点

在图 5-23 中，因为某些轨道首先逼近原点而后又改变方向并远离原点，所以称原点为动态系统的**鞍点**(saddle point). 当矩阵 A 同时有正的和负的本征值时，鞍点就会出现. 排斥最快的方向是沿过 v_1 和 $\mathbf{0}$ 的直线，其中 v_1 对应于正本征值. 吸引最快的方向是沿过 v_2 和 $\mathbf{0}$ 的直线，其中 v_2 对应于负本征值.

357

5.7.1 解耦动态系统

下面的讨论将说明：当 A 是 $n\times n$ 矩阵并且有 n 个线性无关的本征向量时，即当 A 可对角化时，对于由 $\boldsymbol{x}'=A\boldsymbol{x}$ 刻画的任一动态系统，都可以根据例题 1 和例题 2 中的方法求它的一个基础解系. 设 A 的特征函数为

$$v_1e^{\lambda_1 t},\cdots,v_ne^{\lambda_n t}$$

其中 $v_1,\cdots,v_n$ 为线性无关的本征向量. 令 $P=[v_1\ \ \cdots\ \ v_n]$，并且令 D 是对角线元素为 $\lambda_1,\cdots,\lambda_n$ 的对角矩阵，从而 $A=PDP^{-1}$. 现在作变量替换，定义一个新的函数为

$$\boldsymbol{y}(t)=P^{-1}\boldsymbol{x}(t)\text{，或等价地，}\boldsymbol{x}(t)=P\boldsymbol{y}(t)$$

方程 $\boldsymbol{x}(t)=P\boldsymbol{y}(t)$ 说明 $\boldsymbol{y}(t)$ 是 $\boldsymbol{x}(t)$ 关于本征向量基的坐标向量. 用 $P\boldsymbol{y}$ 代替方程 $\boldsymbol{x}'=A\boldsymbol{x}$ 中的 $\boldsymbol{x}$ 得

$$\frac{d}{dt}(P\boldsymbol{y})=A(P\boldsymbol{y})=(PDP^{-1})P\boldsymbol{y}=PD\boldsymbol{y} \tag{5}$$

因为 P 是一个常数矩阵，所以(5)的左边等于 $P\boldsymbol{y}'$. (5) 两边同时左乘 P^{-1} 得 $\boldsymbol{y}'=D\boldsymbol{y}$，或

$$\begin{bmatrix} y'_1(t) \\ y'_2(t) \\ \vdots \\ y'_n(t) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{bmatrix}$$

从 $\boldsymbol{x}$ 到 $\boldsymbol{y}$ 的变量替换实现了对微分方程组的解耦，因为每个数量函数 y_k 的导数仅依赖于 y_k（回忆5.6节中也有类似的变量替换）．因为 $y'_1 = \lambda y_1$，所以 $y_1 = c_1 e^{\lambda_1 t}$，类似地可以得到 $y_2, \cdots, y_n$ 的公式．从而

$$\boldsymbol{y}(t) = \begin{bmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{bmatrix}, \text{其中} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \boldsymbol{y}(0) = \boldsymbol{P}^{-1}\boldsymbol{x}(0) = \boldsymbol{P}^{-1}\boldsymbol{x}_0$$

为了得到原始系统的通解 $\boldsymbol{x}$，计算

$$\boldsymbol{x}(t) = \boldsymbol{P}\boldsymbol{y}(t) = [\boldsymbol{v}_1 \cdots \boldsymbol{v}_n]\boldsymbol{y}(t) = c_1 \boldsymbol{v}_1 e^{\lambda_1 t} + \cdots + c_n \boldsymbol{v}_n e^{\lambda_n t}$$

358 这就如同例题1所构造的特征函数的展开．

5.7.2　复本征值

在下一道例题中，实矩阵 $\boldsymbol{A}$ 有一对复本征值 λ 和 $\bar{\lambda}$，对应的复本征向量为 $\boldsymbol{v}$ 和 $\bar{\boldsymbol{v}}$．回忆5.5节，对于实矩阵，复本征值和对应的复本征向量以共轭对的形式成对出现．所以 $\boldsymbol{x}' = \boldsymbol{A}\boldsymbol{x}$ 的两个解为

$$\boldsymbol{x}_1(t) = \boldsymbol{v}e^{\lambda t} \text{和} \boldsymbol{x}_2(t) = \bar{\boldsymbol{v}}e^{\bar{\lambda} t} \tag{6}$$

利用指数函数的幂级数表示，可以证明 $\boldsymbol{x}_2(t) = \overline{\boldsymbol{x}_1(t)}$．虽然复特征函数 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 对于某些计算（特别是在电气工程方面）来说十分方便，但在许多场合实函数却更适用．庆幸的是，$\boldsymbol{x}_1$ 的实部和虚部是 $\boldsymbol{x}' = \boldsymbol{A}\boldsymbol{x}$ 的（实）解，因为它们是(6)中解的线性组合：

$$\text{Re}(\boldsymbol{v}e^{\lambda t}) = \frac{1}{2}[\boldsymbol{x}_1(t) + \overline{\boldsymbol{x}_1(t)}], \text{Im}(\boldsymbol{v}e^{\lambda t}) = \frac{1}{2i}[\boldsymbol{x}_1(t) - \overline{\boldsymbol{x}_1(t)}]$$

为了理解 $\text{Re}(\boldsymbol{v}e^{\lambda t})$ 的本质，回忆微积分中对于任意的数 x，指数函数 e^x 可以由下面的幂级数来计算：

$$e^x = 1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + \cdots$$

当 λ 为复数时，可以用这个级数来定义 $e^{\lambda t}$：

$$e^{\lambda t} = 1 + (\lambda t) + \frac{1}{2!}(\lambda t)^2 + \cdots + \frac{1}{n!}(\lambda t)^n + \cdots$$

记 $\lambda = a + bi$（a, b 为实数），并且由余弦和正弦函数的幂级数得：

$$e^{(a+bi)t} = e^{at} \cdot e^{ibt} = e^{at}(\cos bt + i \sin bt) \tag{7}$$

于是

$$\begin{aligned} \boldsymbol{v}e^{\lambda t} &= (\text{Re}\,\boldsymbol{v} + i\,\text{Im}\,\boldsymbol{v}) \cdot e^{at}(\cos bt + i \sin bt) \\ &= [(\text{Re}\,\boldsymbol{v})\cos bt - (\text{Im}\,\boldsymbol{v})\sin bt]e^{at} \\ &\quad + i[(\text{Re}\,\boldsymbol{v})\sin bt + (\text{Im}\,\boldsymbol{v})\cos bt]e^{at} \end{aligned}$$

所以 $\boldsymbol{x}' = \boldsymbol{A}\boldsymbol{x}$ 的两个实解为

$$\boxed{\begin{aligned} y_1(t) &= \text{Re}\, x_1(t) = [(\text{Re}\, v)\cos bt - (\text{Im}\, v)\sin bt]e^{at} \\ y_2(t) &= \text{Im}\, x_1(t) = [(\text{Re}\, v)\sin bt + (\text{Im}\, v)\cos bt]e^{at} \end{aligned}}$$

359

可以证明 y_1 和 y_2 是线性无关函数(当 $b \neq 0$ 时).[1]

【例题3】 图5-24中的电路可以由下面的方程描述

$$\begin{bmatrix} i'_L \\ v'_C \end{bmatrix} = \begin{bmatrix} -R_2/L & -1/L \\ 1/C & -1/(R_1C) \end{bmatrix} \begin{bmatrix} i_L \\ v_C \end{bmatrix}$$

其中 i_L 为通过感应体 L 的电流，v_C 为电容器 C 所带的电压. 设 R_1 为5欧姆，R_2 为0.8欧姆，C 为1法拉，L 为0.4亨利. 如果通过感应器的初始电流为3安培，并且电容器的初始电压为3伏特，试求 i_L 和 v_C 的公式.

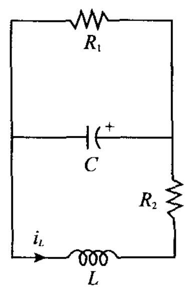


图 5-24

解： 根据已知数据，令 $A = \begin{bmatrix} -2 & -2.5 \\ 10 & -2 \end{bmatrix}$, $x_0 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$. 由5.5节中的方法可以得到本征值 $\lambda = -2+5i$，以及对应的本征向量 $v_1 = \begin{bmatrix} i \\ 2 \end{bmatrix}$. $x' = Ax$ 的复解为

$$x_1(t) = \begin{bmatrix} i \\ 2 \end{bmatrix} e^{(-2+5i)t} \text{和}\ x_2(t) = \begin{bmatrix} -i \\ 2 \end{bmatrix} e^{(-2-5i)t}$$

的线性组合. 下面，根据(7)得

$$x_1(t) = \begin{bmatrix} i \\ 2 \end{bmatrix} e^{-2t}(\cos 5t + i \sin 5t)$$

x_1 的实部和虚部为实解：

$$y_1(t) = \begin{bmatrix} -\sin 5t \\ 2\cos 5t \end{bmatrix} e^{-2t}, \quad y_2(t) = \begin{bmatrix} \cos 5t \\ 2\sin 5t \end{bmatrix} e^{-2t}$$

因为 y_1 和 y_2 是线性无关函数，所以它们构成了 $x' = Ax$ 的全部解所组成的二维实向量空间的一组基. 从而通解为

$$x(t) = c_1 \begin{bmatrix} -\sin 5t \\ 2\cos 5t \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} \cos 5t \\ 2\sin 5t \end{bmatrix} e^{-2t}$$

若要满足 $x(0) = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$，则需 $c_1 \begin{bmatrix} 0 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$，得 $c_1 = 1.5$, $c_2 = 3$. 因此

$$x(t) = 1.5 \begin{bmatrix} -\sin 5t \\ 2\cos 5t \end{bmatrix} e^{-2t} + 3 \begin{bmatrix} \cos 5t \\ 2\sin 5t \end{bmatrix} e^{-2t}$$

---

1. 因为 $x_2(t)$ 是 $x_1(t)$ 的复共轭，所以 $x_2(t)$ 的实部和虚部分别为 $y_1(t)$ 和 $-y_2(t)$. 从而由 $x_1(t)$ 或 $x_2(t)$ 都可以得到 $x' = Ax$ 的两个实线性无关解，但不能同时使用 $x_1(t)$ 和 $x_2(t)$.

或

$$\begin{bmatrix} i_L(t) \\ v_C(t) \end{bmatrix} = \begin{bmatrix} -1.5\sin 5t + 3\cos 5t \\ 3\cos 5t + 6\sin 5t \end{bmatrix} e^{-2t}$$

见图5-25. ■

在图5-25中，原点称为动态系统的**螺线极点**(spiral point). 旋转是由复本征值中出现的正弦和余弦函数导致的. 轨道向内旋转是因为因式 e^{-2t} 趋于零. 其中 -2 是例题3中本征值的实部. 当 $\boldsymbol{A}$ 的复本征值的实部为正时，轨道向外旋转. 如果实部为零，则轨道为环绕原点的椭圆.

360

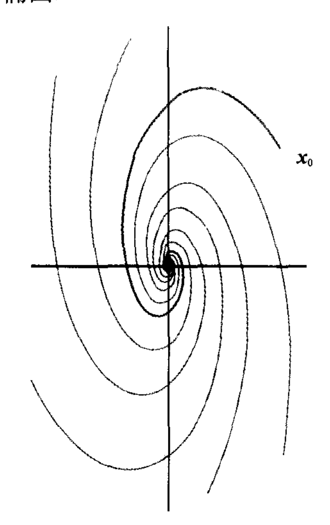


图5-25　原点为螺线极点

基础练习

一个实 3×3 矩阵 $\boldsymbol{A}$ 的本征值为 -0.5，$0.2+0.3i$ 和 $0.2-0.3i$，对应的本征向量为

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix},\quad \boldsymbol{v}_2 = \begin{bmatrix} 1+2i \\ 4i \\ 2 \end{bmatrix},\quad \boldsymbol{v}_3 = \begin{bmatrix} 1-2i \\ -4i \\ 2 \end{bmatrix}$$

1. 使用复矩阵，$\boldsymbol{A}$ 可对角化为 $\boldsymbol{A}=\boldsymbol{PDP}^{-1}$ 吗？
2. 利用复特征函数写出 $\boldsymbol{x}'=\boldsymbol{Ax}$ 的通解，然后求出实通解.
3. 描述典型轨道的形状.

习题 5.7

1. 质点在平面力场中运动的位置向量 $\boldsymbol{x}$ 满足 $\boldsymbol{x}'=\boldsymbol{Ax}$. $\boldsymbol{A}$ 为 2×2 矩阵，本征值为4和 -2，对应的本征向量为 $\boldsymbol{v}_1=\begin{bmatrix} -3 \\ 1 \end{bmatrix}$ 和 $\boldsymbol{v}_2=\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ 设 $\boldsymbol{x}(0)=\begin{bmatrix} -6 \\ 1 \end{bmatrix}$，试求质点在时刻 t 的位置.
2. 设 $\boldsymbol{A}$ 为 2×2 矩阵，本征值为 -3 和 -1，对应的本征向量为 $\boldsymbol{v}_1=\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ 和 $\boldsymbol{v}_2=\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. 令 $\boldsymbol{x}(t)$

为质点在时刻 t 的位置. 求解初值问题 $\boldsymbol{x}'=\boldsymbol{A}\boldsymbol{x}$, $\boldsymbol{x}(0)=\begin{bmatrix}2\\3\end{bmatrix}$.

在习题3~6中，对 $t\geqslant 0$ 求解初值问题 $\boldsymbol{x}'=\boldsymbol{A}\boldsymbol{x},\boldsymbol{x}(\boldsymbol{0})=(3,2)$. 并将原点的性质按照动态系统的吸引子、排斥子或鞍点进行分类，其中的动态系统由 $\boldsymbol{x}'=\boldsymbol{A}\boldsymbol{x}$ 描述. 求出吸引和/或排斥最快的方向. 当原点为鞍点时，绘出典型轨道的草图.

3. $\boldsymbol{A}=\begin{bmatrix}2&3\\-1&-2\end{bmatrix}$　4. $\boldsymbol{A}=\begin{bmatrix}-2&-5\\1&4\end{bmatrix}$　5. $\boldsymbol{A}=\begin{bmatrix}7&-1\\3&3\end{bmatrix}$　6. $\boldsymbol{A}=\begin{bmatrix}1&-2\\3&-4\end{bmatrix}$

在习题7和8中，做变量替换使之解耦方程 $\boldsymbol{x}'=\boldsymbol{A}\boldsymbol{x}$. 指定 $\boldsymbol{P}$ 和 $\boldsymbol{D}$，写出方程 $\boldsymbol{x}(t)=\boldsymbol{P}\boldsymbol{y}(t)$ 并给出解耦系统 $\boldsymbol{y}'=\boldsymbol{D}\boldsymbol{y}$ 的推导过程.

7. $\boldsymbol{A}$ 同习题5　8. $\boldsymbol{A}$ 同习题6

在习题9~18中，用复特征函数构造 $\boldsymbol{x}'=\boldsymbol{A}\boldsymbol{x}$ 的通解，然后写出实通解. 描述典型轨道的形状.

9. $\boldsymbol{A}=\begin{bmatrix}-3&2\\-1&-1\end{bmatrix}$　10. $\boldsymbol{A}=\begin{bmatrix}3&1\\-2&1\end{bmatrix}$　11. $\boldsymbol{A}=\begin{bmatrix}-3&-9\\2&3\end{bmatrix}$

12. $\boldsymbol{A}=\begin{bmatrix}-7&10\\-4&5\end{bmatrix}$　13. $\boldsymbol{A}=\begin{bmatrix}4&-3\\6&-2\end{bmatrix}$　14. $\boldsymbol{A}=\begin{bmatrix}-2&1\\-8&2\end{bmatrix}$

15. [M] $\boldsymbol{A}=\begin{bmatrix}-8&-12&-6\\2&1&2\\7&12&5\end{bmatrix}$　16. [M] $\boldsymbol{A}=\begin{bmatrix}-6&-11&16\\2&5&-4\\-4&-5&10\end{bmatrix}$

17. [M] $\boldsymbol{A}=\begin{bmatrix}30&64&23\\-11&-23&-9\\6&15&4\end{bmatrix}$　18. [M] $\boldsymbol{A}=\begin{bmatrix}53&-30&-2\\90&-52&-3\\20&-10&2\end{bmatrix}$

361

19. [M] 对于例题1中的电路，设 $R_1=1/5$ 欧姆，$R_2=1/3$ 欧姆，$C_1=4$ 法拉，$C_2=3$ 法拉，并且每个电容器的初始电压为3伏特，试求电压 v_1 和 v_2(看成 t 的函数)的公式.
20. [M] 对于例题1中的电路，设 $R_1=1/15$ 欧姆，$R_2=1/3$ 欧姆，$C_1=9$ 法拉，$C_2=2$ 法拉，并且每个电容器的初始电压为3伏特，试求电压 v_1 和 v_2 的公式.
21. [M] 对于例题3中的电路，设 $R_1=1$ 欧姆，$R_2=0.125$ 欧姆，$C_1=0.2$ 法拉，$L=0.125$ 亨利，初始电流为0安培，初始电压为15伏特，试求电流 I_L 和电压 v_C 的公式.
22. [M]右图中的电路可由下面的方程描述

$$\begin{bmatrix}i'_L\\v'_C\end{bmatrix}=\begin{bmatrix}0&1/L\\-1/C&-1/(RC)\end{bmatrix}\begin{bmatrix}i_L\\v_C\end{bmatrix}$$

R
\+
C
L

其中 i_L 为通过感应体 L 的电流，v_C 为电容器 C 的电压. 设 $R=0.5$ 欧姆，$C=2.5$ 法拉，$L=0.5$ 亨利，初始电流为0安培，初始电压为12伏特，试求 i_L 和 v_C 的公式.

基础练习答案

1. 可以. 因为3×3矩阵有三个不同的本征值，所以它可对角化. 当有复数量时，5.1节定理2和5.3节定理5仍有效(其证明本质上与实数量的情形是一样的).
2. 通解形如

$$x(t)=c_1\begin{bmatrix}1\\-2\\1\end{bmatrix}e^{-0.5t}+c_2\begin{bmatrix}1+2i\\4i\\2\end{bmatrix}e^{(0.2+0.3i)t}+c_3\begin{bmatrix}1-2i\\-4i\\2\end{bmatrix}e^{(0.2-0.3i)t}$$

这里的数量 c_1,c_2,c_3 可以为任何复数. $x(t)$ 中的第一项是实的. 另外两个实解可以由 $x(t)$ 中的第二项的实部和虚部求得：

$$\begin{bmatrix}1+2i\\4i\\2\end{bmatrix}e^{0.2t}(\cos 0.3t+i\sin 0.3t)$$

实通解形如：

$$c_1\begin{bmatrix}1\\-2\\1\end{bmatrix}e^{-0.5t}+c_2\begin{bmatrix}\cos 0.3t-2\sin 0.3t\\-4\sin 0.3t\\2\cos 0.3t\end{bmatrix}e^{0.2t}+c_3\begin{bmatrix}\sin 0.3t+2\cos 0.3t\\4\cos 0.3t\\2\sin 0.3t\end{bmatrix}e^{0.2t}$$

其中 c_1,c_2,c_3 为实数.

3. 因为 $e^{-0.5t}$ 有负指数因数，所以任一 $c_2=c_3=0$ 的解都被吸引到原点. 其他的解因为有无界的分量，所以轨道向外盘旋.

362

当心不要将本题错认为5.6节中的一道题目. 在5.6节中，吸引到 **0** 的条件是本征值的绝对值小于1，以保证 $|\lambda|^k\to 0$. 而这里本征向量的实部必须为负的，以保证 $e^{\lambda t}\to 0$.

5.8 本征值的迭代估计

在线性代数的科学应用中，本征值很少能精确地求得，不过，庆幸的是通常都能找到令人满意的近似值. 事实上，有些应用只需要对最大的本征值进行粗略的近似. 对于这种情况，下面介绍的第一个算法非常有效. 同时，它为一个更有效的方法奠定了基础，这个更有效的方法可以给出其他本征值的快速估计.

5.8.1 幂法

幂法适用于 $n\times n$ 矩阵 A，其中 A 有**严格占优本征值**(strictly dominant eigenvalue) λ_1，即 λ_1 的绝对值必须大于所有其他本征值的绝对值. 在这种情形下，幂法生成一个逼近 λ_1 的数量序列和一个逼近其对应本征向量的向量序列. 方法的背景是在5.6节开始时提到的本征向量分解.

简单起见，设 A 可对角化，本征向量 $v_1,\cdots,v_n$ 为 $\mathbf{R}^n$ 的一组基，对应的本征值 $\lambda_1,\cdots,\lambda_n$ 依次递减，λ_1 为严格占优本征值. 即

$$|\lambda_1|>|\lambda_2|\geqslant|\lambda_3|\geqslant\cdots\geqslant|\lambda_n| \tag{1}$$

（$>$ 处标注：严格大于）

如5.6节中(2)所示，如果 $\mathbf{R}^n$ 中向量 x 可以写为 $x=c_1v_1+\cdots+c_nv_n$，则

$$A^kx=c_1(\lambda_1)^kv_1+c_2(\lambda_2)^kv_2+\cdots+c_n(\lambda_n)^kv_n\ (k=1,2,\cdots)$$

设 $c_1\neq 0$，然后除以 $(\lambda_1)^k$，得

$$\frac{1}{(\lambda_1)^k}A^kx=c_1v_1+c_2\left(\frac{\lambda_2}{\lambda_1}\right)^kv_2+\cdots+c_n\left(\frac{\lambda_n}{\lambda_1}\right)^kv_n\ (k=1,2,\cdots) \tag{2}$$

由(1)知，分数 $\lambda_2/\lambda_1,\cdots,\lambda_n/\lambda_1$ 的绝对值均小于1，因而它们的幂趋于零. 所以

$$(\lambda_1)^{-k}A^kx\to c_1v_1,\text{当 } k\to\infty \text{ 时} \tag{3}$$

因此,对于足够大的 k,A^kx 的数量倍数确定的方向与本征向量 c_1v_1 几乎一致.由于数乘一个正数并不改变向量的方向,所以若 $c_1\neq 0$,则 A^kx 和 v_1 或 $-v_1$ 的方向几乎一致.

【例题1】 设 $A=\begin{bmatrix}1.8 & 0.8\\0.2 & 1.2\end{bmatrix}$,$v_1=\begin{bmatrix}4\\1\end{bmatrix}$,$x=\begin{bmatrix}-0.5\\1\end{bmatrix}$.则 A 的本征值为 2 和 1,$\lambda_1=2$ 对应的本征空间为通过 $\mathbf{0}$ 和 v_1 的直线.对 $k=0,\cdots,8$,计算 A^kx 并构造通过 $\mathbf{0}$ 和 A^kx 的直线. 随着 k 的增加有什么变化?

363

解: 前三项的计算如下

$$Ax=\begin{bmatrix}1.8 & 0.8\\0.2 & 1.2\end{bmatrix}\begin{bmatrix}-0.5\\1\end{bmatrix}=\begin{bmatrix}-0.1\\1.1\end{bmatrix}$$

$$A^2x=A(Ax)=\begin{bmatrix}1.8 & 0.8\\0.2 & 1.2\end{bmatrix}\begin{bmatrix}-0.1\\1.1\end{bmatrix}=\begin{bmatrix}0.7\\1.3\end{bmatrix}$$

$$A^3x=A(A^2x)=\begin{bmatrix}1.8 & 0.8\\0.2 & 1.2\end{bmatrix}\begin{bmatrix}0.7\\1.3\end{bmatrix}=\begin{bmatrix}2.3\\1.7\end{bmatrix}$$

类似的计算得到表 5-1.

表 5-1 向量的迭代

k	0	1	2	3	4	5	6	7	8
A^kx	$\begin{bmatrix}-0.5\\1\end{bmatrix}$	$\begin{bmatrix}-0.1\\1.1\end{bmatrix}$	$\begin{bmatrix}0.7\\1.3\end{bmatrix}$	$\begin{bmatrix}2.3\\1.7\end{bmatrix}$	$\begin{bmatrix}5.5\\2.5\end{bmatrix}$	$\begin{bmatrix}11.9\\4.1\end{bmatrix}$	$\begin{bmatrix}24.7\\7.3\end{bmatrix}$	$\begin{bmatrix}50.3\\13.7\end{bmatrix}$	$\begin{bmatrix}101.5\\26.5\end{bmatrix}$

向量 $x,Ax,\cdots,A^4x$ 如图 5-26 所示. 其他的向量因太长而不能在本图中显示. 尽管如此,仍画出线段来表示那些向量的方向. 事实上,我们关心的也是向量的方向而不是向量本身. 这些直线似乎趋于一条直线,该直线代表由 v_1 生成的本征空间. 更精确的说,当 $k\to\infty$ 时,由 A^kx 确定的直线(子空间)与由 v_1 确定的直线(本征空间)之间的夹角趋于零. ■

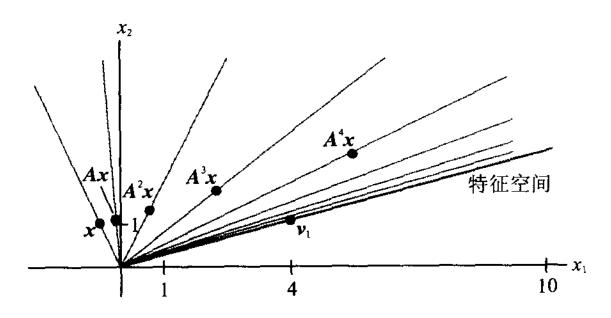


图 5-26 x, Ax, A^2x, ⋯, A^7x 所确定的方向

倘若 $c_1\neq 0$, (3) 中的向量 $(\lambda_1)^{-k}A^kx$ 经过缩放收敛于 c_1v_1. 但由于不知道 λ_1, 我们无法对 A^kx 进行同样的缩放, 不过可以对每个 A^kx 进行缩放使其最大的元素为 1. 所

得到的序列$\{x_k\}$收敛到v_1的某个倍数，其中最大的元素为1．图5-27绘出了例题1缩放后的序列．本征值λ_1也可以根据序列$\{x_k\}$做出估计．当x_k逼近λ_1所对应的本征向量时，向量Ax_k逼近$\lambda_1 x_k$，Ax_k中的每个元素都近似于x_k中对应元素的λ_1倍．因为x_k中最大的元素为1，所以Ax_k中最大的元素近似于1．（以上结论的详细证明在此略过．）

364

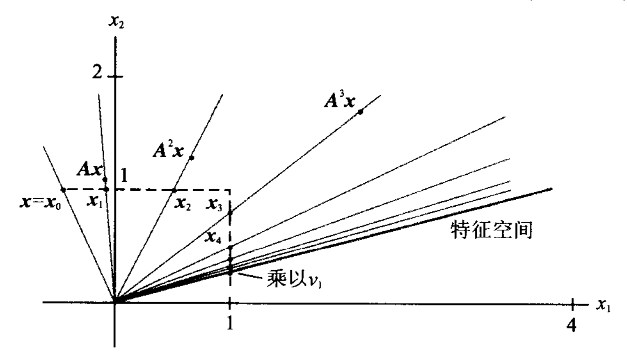


图5-27　x，Ax，A^2x，…，A^7x的数量倍数

估计严格占优本征值的幂法

1. 选定初始向量x_0，其最大的元素为1.
2. 对于$k=0,1,\cdots$，
   a. 计算Ax_k.
   b. 令μ_k为Ax_k中绝对值最大的元素.
   c. 计算$x_{k+1}=(1/\mu_k)Ax_k$.
3. 对于所有的x_0，序列$\{\mu_k\}$趋于占优本征值，并且序列$\{x_k\}$趋于对应的本征向量.

【例题2】　设$A=\begin{bmatrix}6 & 5\\ 1 & 2\end{bmatrix}$，$x_0=\begin{bmatrix}0\\ 1\end{bmatrix}$．取$k=5$，用幂法估计$A$的占优本征值和对应的本征向量.

解：本题及下一道例题的计算由MATLAB完成，计算结果精确到小数点后16位，不过这里我们只给出前几位．首先，计算Ax_0并求Ax_0中最大的元素μ_0：

$$Ax_0=\begin{bmatrix}6 & 5\\ 1 & 2\end{bmatrix}\begin{bmatrix}0\\ 1\end{bmatrix}=\begin{bmatrix}5\\ 2\end{bmatrix},\quad \mu_0=5$$

Ax_0乘以$1/\mu_0$得到x_1，计算Ax_1，并求Ax_1中最大的元素：

$$x_1=\frac{1}{\mu_0}Ax_0=\frac{1}{5}\begin{bmatrix}5\\ 2\end{bmatrix}=\begin{bmatrix}1\\ 0.4\end{bmatrix}$$

$$Ax_1=\begin{bmatrix}6 & 5\\ 1 & 2\end{bmatrix}\begin{bmatrix}1\\ 0.4\end{bmatrix}=\begin{bmatrix}8\\ 1.8\end{bmatrix},\quad \mu_1=8$$

Ax_1乘以 $1/\mu_1$ 得到 x_2，计算 Ax_2，并求 Ax_2中最大的元素：

$$x_2=\frac{1}{\mu_1}Ax_1=\frac{1}{8}\begin{bmatrix}8\\1.8\end{bmatrix}=\begin{bmatrix}1\\0.225\end{bmatrix}$$

$$Ax_2=\begin{bmatrix}6&5\\1&2\end{bmatrix}\begin{bmatrix}1\\0.225\end{bmatrix}=\begin{bmatrix}7.125\\1.450\end{bmatrix},\quad \mu_2=7.125$$

Ax_2乘以 $1/\mu_2$ 得到 x_3，等等. 表 5-2 中列出了 MATLAB 前五次迭代的结果.

表 5-2 例题 2 的幂法

k	0	1	2	3	4	5
x_k	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0.4\end{bmatrix}$	$\begin{bmatrix}1\\0.225\end{bmatrix}$	$\begin{bmatrix}1\\0.2035\end{bmatrix}$	$\begin{bmatrix}1\\0.2005\end{bmatrix}$	$\begin{bmatrix}1\\0.20007\end{bmatrix}$
Ax_k	$\begin{bmatrix}5\\2\end{bmatrix}$	$\begin{bmatrix}8\\1.8\end{bmatrix}$	$\begin{bmatrix}7.125\\1.450\end{bmatrix}$	$\begin{bmatrix}7.0175\\1.4070\end{bmatrix}$	$\begin{bmatrix}7.0025\\1.4010\end{bmatrix}$	$\begin{bmatrix}7.00036\\1.40014\end{bmatrix}$
μ_k	5	8	7.125	7.0175	7.0025	7.00036

表 5-2 中的数据表明$\{x_k\}$趋于(1,0.2)且$\{\mu_k\}$趋于 7. 这样就有（1，0.2）为本征向量且 7 为占优本征值. 这很容易验证，只需计算

$$A\begin{bmatrix}1\\0.2\end{bmatrix}=\begin{bmatrix}6&5\\1&2\end{bmatrix}\begin{bmatrix}1\\0.2\end{bmatrix}=\begin{bmatrix}7\\1.4\end{bmatrix}=7\begin{bmatrix}1\\0.2\end{bmatrix}$$ ■

例题 2 中的序列$\{\mu_k\}$之所以能迅速收敛到 $\lambda_1=7$ 是因为 A 的第二个本征值很小（事实上，$\lambda_2=1$）. 一般地，收敛的速率依赖于比值$|\lambda_2/\lambda_1|$，这是因为当缩放 A^kx 来估计 c_1v_1时，(2) 中的向量 $c_2(\lambda_2/\lambda_1)^kv_2$ 是误差的主要来源（其他分式 λ_j/λ_1 较小）. 如果$|\lambda_2/\lambda_1|$逼近 1，则$\{\mu_k\}$和$\{x_k\}$的收敛会非常缓慢，或许其他逼近方法会更适合.

对于幂法，选定的初始向量x 可能在 v_1方向上没有分量（当 $c_1=0$ 时）. 但计算机在计算 x_k时的舍入误差可能会产生一个向量，该向量在 v_1方向上至少有一个小的分量. 如果这种情况出现，则 x_k将开始收敛到 v_1的倍数.

5.8.2 逆幂法

若已知本征值 λ 的一个较好的估计 α，则逆幂法可以逼近任何本征值. 此时我们令 $B=(A-\alpha I)^{-1}$并对 B 使用幂法. 可以证明，如果 A 的本征值为 $\lambda_1,\cdots,\lambda_n$，则 B 的本征值为 [366]

$$\frac{1}{\lambda_1-\alpha},\frac{1}{\lambda_2-\alpha},\cdots,\frac{1}{\lambda_n-\alpha}$$

且对应的本征向量和 A 的本征向量一样（见习题 15 和 16）.

例如，假设和 A 的其他本征值相比，α 更接近 λ_2. 则 $1/(\lambda_2-\alpha)$将为 B 的严格占优本征值. 如果 α 真的近似于 λ_2，则 $1/(\lambda_2-\alpha)$远大于 B 的其他本征值，并且对于任意的 x_0，逆幂法都可以生成 λ_2 的一个快速逼近. 下面的算法给出了细节.

估计 A 的本征值 λ 的逆幂法

1. 选定一个初始估计值 α，使其尽量接近 λ.

2. 选定一个初始向量，其最大的元素为1.
3. 对于 $k=0,1,\cdots$,
   a. 求解 $\boldsymbol{y}_k$ 的方程 $(\boldsymbol{A}-\alpha\boldsymbol{I})\ \boldsymbol{y}_k=\boldsymbol{x}_k$.
   b. 令 μ_k 为 $\boldsymbol{y}_k$ 中绝对值最大的元素.
   c. 计算 $\nu_k=\alpha+(1/\mu_k)$.
   d. 计算 $\boldsymbol{x}_{k+1}=(1/\mu_k)\boldsymbol{y}_k$.
4. 对于几乎所有的 $\boldsymbol{x}_0$，序列 $\{\nu_k\}$ 趋于 $\boldsymbol{A}$ 的本征值 λ，并且序列 $\{\boldsymbol{x}_k\}$ 趋于对应的本征向量.

注意到 $\boldsymbol{B}$，或者 $(\boldsymbol{A}-\alpha\boldsymbol{I})^{-1}$，并没有出现在算法之中. 与其计算 $(\boldsymbol{A}-\alpha\boldsymbol{I})^{-1}\boldsymbol{x}_k$ 得到序列中的下一个向量，不如求解 $\boldsymbol{y}_k$ 的方程 $(\boldsymbol{A}-\alpha\boldsymbol{I})\boldsymbol{y}_k=\boldsymbol{x}_k$（然后缩放 $\boldsymbol{y}_k$ 以生成 $\boldsymbol{x}_{k+1}$）. 由于对每个 k 都要求解 $\boldsymbol{y}_k$ 的方程，使用 $\boldsymbol{A}-\alpha\boldsymbol{I}$ 的LU分解可以加快计算.

【例题3】　在某些应用中，需要知道矩阵 $\boldsymbol{A}$ 的最小的本征值以及本征值的粗略估计. 设21，3.3和1.9是下面矩阵 $\boldsymbol{A}$ 的本征值的估计. 试求最小本征值，精确到六位小数.

$$\boldsymbol{A}=\begin{bmatrix}10 & -8 & -4\\ -8 & 13 & 4\\ -4 & 5 & 4\end{bmatrix}$$

解：最小的两个本征值比较接近，所以我们对 $\boldsymbol{A}-1.9\boldsymbol{I}$ 使用逆幂法. MATLAB计算的结果如表5-3所示. 这里 $\boldsymbol{x}_0$ 是任意选取的，$\boldsymbol{y}_k=(\boldsymbol{A}-1.9\boldsymbol{I})^{-1}\boldsymbol{x}_k,\mu_k$ 为 $\boldsymbol{y}_k$ 中最大的元素，$\nu_k=1.9+1/\mu_k$，$\boldsymbol{x}_{k+1}=(1/\mu_k)\boldsymbol{y}_k$. 结果表明，初始本征值的估计非常好，并且逆幂序列收敛地也非常快. 准确的最小本征值为2.

367

表5-3　逆幂法

k	0	1	2	3	4
$\boldsymbol{x}_k$	$\begin{bmatrix}1\\1\\1\end{bmatrix}$	$\begin{bmatrix}0.5736\\0.0646\\1\end{bmatrix}$	$\begin{bmatrix}0.5054\\0.0045\\1\end{bmatrix}$	$\begin{bmatrix}0.5004\\0.0003\\1\end{bmatrix}$	$\begin{bmatrix}0.50003\\0.00002\\1\end{bmatrix}$
$\boldsymbol{y}_k$	$\begin{bmatrix}4.45\\0.50\\7.76\end{bmatrix}$	$\begin{bmatrix}5.0131\\0.0442\\9.9197\end{bmatrix}$	$\begin{bmatrix}5.0012\\0.0031\\9.9949\end{bmatrix}$	$\begin{bmatrix}5.0001\\0.0002\\9.9996\end{bmatrix}$	$\begin{bmatrix}5.000006\\0.000015\\9.999975\end{bmatrix}$
μ_k	7.76	9.9197	9.9949	9.9996	9.999975
ν_k	2.03	2.0008	2.00005	2.000004	2.0000002

■

如果得不到对矩阵最小本征值的估计，在逆幂法中可以简单地取 $\alpha=0$. 当最小的本征值比其他本征值更接近零时，上述 α 的选择很有效.

对于许多简单的情况，本节介绍的两个算法很实用，并且它们也使我们对本征值估计问题有了一定的了解. 一个更加有效且被广泛应用的方法是QR算法. 例如，它是MATLAB中的命令 `eig(A)` 的核心，这个命令可以快速计算 $\boldsymbol{A}$ 的本征值和本征向量. 在5.2节的习题中曾给出了QR算法的简要描述. 大部分现代数值分析教材中也都有进一步的讨论.

基础练习

怎样判定一个已知向量 $\boldsymbol{x}_0$ 是矩阵 $\boldsymbol{A}$ 的本征向量的一个好的近似？如果它是，怎样估计对应的本征值？用下面的矩阵和向量验证你的方法.

$$\boldsymbol{A}=\begin{bmatrix}5 & 8 & 4\\ 8 & 3 & -1\\ 4 & -1 & 2\end{bmatrix}\text{和}\ \boldsymbol{x}=\begin{bmatrix}1.0\\ -4.3\\ 8.1\end{bmatrix}$$

习题 5.8

在习题 1 ~ 4 中，矩阵 $\boldsymbol{A}$ 后面是由幂法生成的序列 $\{\boldsymbol{x}_k\}$. 根据这些数据估计 $\boldsymbol{A}$ 最大的本征值以及对应的本征向量.

1. $\boldsymbol{A}=\begin{bmatrix}4 & 3\\ 1 & 2\end{bmatrix}$; $\begin{bmatrix}1\\ 0\end{bmatrix}$, $\begin{bmatrix}1\\ 0.25\end{bmatrix}$, $\begin{bmatrix}1\\ 0.3158\end{bmatrix}$, $\begin{bmatrix}1\\ 0.3298\end{bmatrix}$, $\begin{bmatrix}1\\ 0.3326\end{bmatrix}$

368

2. $\boldsymbol{A}=\begin{bmatrix}1.8 & -0.8\\ -3.2 & 4.2\end{bmatrix}$; $\begin{bmatrix}1\\ 0\end{bmatrix}$, $\begin{bmatrix}-0.5625\\ 1\end{bmatrix}$, $\begin{bmatrix}-0.3021\\ 1\end{bmatrix}$, $\begin{bmatrix}-0.2601\\ 1\end{bmatrix}$, $\begin{bmatrix}-0.2520\\ 1\end{bmatrix}$

3. $\boldsymbol{A}=\begin{bmatrix}0.5 & 0.2\\ 0.4 & 0.7\end{bmatrix}$; $\begin{bmatrix}1\\ 0\end{bmatrix}$, $\begin{bmatrix}1\\ 0.8\end{bmatrix}$, $\begin{bmatrix}0.6875\\ 1\end{bmatrix}$, $\begin{bmatrix}0.5577\\ 1\end{bmatrix}$, $\begin{bmatrix}0.5188\\ 1\end{bmatrix}$

4. $\boldsymbol{A}=\begin{bmatrix}4.1 & -6\\ 3 & -4.4\end{bmatrix}$; $\begin{bmatrix}1\\ 1\end{bmatrix}$, $\begin{bmatrix}1\\ 0.7368\end{bmatrix}$, $\begin{bmatrix}1\\ 0.7541\end{bmatrix}$, $\begin{bmatrix}1\\ 0.7490\end{bmatrix}$, $\begin{bmatrix}1\\ 0.7502\end{bmatrix}$

5. 设 $\boldsymbol{A}=\begin{bmatrix}15 & 16\\ -20 & -21\end{bmatrix}$. 向量 $\boldsymbol{x},\cdots,\boldsymbol{A}^5\boldsymbol{x}$ 为 $\begin{bmatrix}1\\ 1\end{bmatrix}$, $\begin{bmatrix}31\\ -41\end{bmatrix}$, $\begin{bmatrix}-191\\ 241\end{bmatrix}$, $\begin{bmatrix}991\\ -1241\end{bmatrix}$, $\begin{bmatrix}-4991\\ 6241\end{bmatrix}$, $\begin{bmatrix}24991\\ -31241\end{bmatrix}$. 试求一个向量，使其第二个元素是 1，且该向量近似于 $\boldsymbol{A}$ 的一个本征向量. 保留四位小数. 检验你的估计值，并给出对 $\boldsymbol{A}$ 的占优本征值的一个估计值.

6. 设 $\boldsymbol{A}=\begin{bmatrix}-2 & -3\\ 6 & 7\end{bmatrix}$. 已知下面序列，重做习题 5.

$$\begin{bmatrix}1\\ 1\end{bmatrix},\ \begin{bmatrix}-5\\ 13\end{bmatrix},\ \begin{bmatrix}-29\\ 61\end{bmatrix},\ \begin{bmatrix}-125\\ 253\end{bmatrix},\ \begin{bmatrix}-509\\ 1021\end{bmatrix},\ \begin{bmatrix}-2045\\ 4093\end{bmatrix}$$

[M]习题 7 ~ 12 需要 MATLAB 或其他计算工具. 在习题 7 和 8 中，已知 $\boldsymbol{x}_0$，对于 $k=1,\cdots,5$ 用幂法列出 $\{\boldsymbol{x}_k\}$ 和 $\{\mu_k\}$. 在习题 9 和 10 中，写出 μ_5 和 μ_6.

7. $\boldsymbol{A}=\begin{bmatrix}6 & 7\\ 8 & 5\end{bmatrix}$, $\boldsymbol{x}_0=\begin{bmatrix}1\\ 0\end{bmatrix}$

8. $\boldsymbol{A}=\begin{bmatrix}2 & 1\\ 4 & 5\end{bmatrix}$, $\boldsymbol{x}_0=\begin{bmatrix}1\\ 0\end{bmatrix}$

9. $\boldsymbol{A}=\begin{bmatrix}8 & 0 & 12\\ 1 & -2 & 1\\ 0 & 3 & 0\end{bmatrix}$, $\boldsymbol{x}_0=\begin{bmatrix}1\\ 0\\ 0\end{bmatrix}$

10. $\boldsymbol{A}=\begin{bmatrix}1 & 2 & -2\\ 1 & 1 & 9\\ 0 & 1 & 9\end{bmatrix}$, $\boldsymbol{x}_0=\begin{bmatrix}1\\ 0\\ 0\end{bmatrix}$

当已知近似的本征向量时，可以估计对应的本征值. 观察到如果 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$，则 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}=$

$x^T(\lambda x)=\lambda(x^Tx)$，并且**瑞利商**(Rayleigh quotient)

$$R(x)=\frac{x^TAx}{x^Tx}$$

等于 λ. 如果 x 近似于 λ 的本征向量，则这个商近似于 λ. 当 A 为对称矩阵($A^T=A$)时，瑞利商 $R(x_k)=(x_k^TAx_k)/(x_k^Tx_k)$ 的精度大概是幂法中缩放因子 μ_k 的两倍. 在习题11和12中，对于 $k=1,\cdots,4$，计算 μ_k 和 $R(x_k)$，并比较其精度.

11. $A=\begin{bmatrix}5 & 2\\ 2 & 2\end{bmatrix}$，$x_0=\begin{bmatrix}1\\ 0\end{bmatrix}$

12. $A=\begin{bmatrix}-3 & 2\\ 2 & 0\end{bmatrix}$，$x_0=\begin{bmatrix}1\\ 0\end{bmatrix}$

在习题13和14中，A 为 3×3 矩阵，对本征值的估计为4，-4 和3.

13. 如果已知接近4和 -4 的本征值的绝对值不相等，可以使用幂法吗？它还有效吗？
14. 假设近似于4和 -4 的本征值的绝对值相等. 怎样得到一个估计近似于4的本征值序列.
15. 设 $Ax=\lambda x$，其中 $x\neq 0$. 令 α 为不同于 A 的本征值的数量，并且令 $B=(A-\alpha I)^{-1}$. 将方程 $Ax=\lambda x$ 的两边都减去 αx，由代数计算证明 $1/(\lambda-\alpha)$ 为 B 的本征值，x 为其对应的本征向量.
16. 设 μ 为习题15中 B 的一个本征值，并且 x 为对应的本征向量，所以 $(A-\alpha I)^{-1}x=\mu x$. 根据这个方程，试求 A 的一个本征值，用 μ 和 α 表示.
17. [M]用逆幂法估计例题3中矩阵 A 的中间本征值，精确到四位小数. 令 $x_0=(1,0,0)$.
18. [M]令 A 为习题9中的矩阵. 用逆幂法估计矩阵 A 在 $\alpha=-1.4$ 左右的本征值，精确到四位小数. 令 $x_0=(1,0,0)$.

[M]在习题19和20中，试求(a)最大的本征值和(b)接近零的本征值. 在两种情形中，都令 $x_0=(1,0,0,0)$ 并精确到四位小数. 给出近似的本征向量.

19. $A=\begin{bmatrix}10 & 7 & 8 & 7\\ 7 & 5 & 6 & 5\\ 8 & 6 & 10 & 9\\ 7 & 5 & 9 & 10\end{bmatrix}$

369

20. $A=\begin{bmatrix}1 & 2 & 3 & 2\\ 2 & 12 & 13 & 11\\ -2 & 3 & 0 & 2\\ 4 & 5 & 7 & 2\end{bmatrix}$

21. 一个常见的误解是如果 A 有严格占优本征值，则对于足够大的 k，向量 A^kx 约等于 A 的本征向量. 对于下面三个矩阵，研究当 $x=(0.5,0.5)$ 时 A^kx 的变化，并尽量给出一般性的结论(对于 2×2 矩阵).

a. $A=\begin{bmatrix}0.8 & 0\\ 0 & 0.2\end{bmatrix}$　b. $A=\begin{bmatrix}1 & 0\\ 0 & 0.8\end{bmatrix}$　c. $A=\begin{bmatrix}8 & 0\\ 0 & 2\end{bmatrix}$

基础练习答案

对于已知的 A 和 x

$$Ax=\begin{bmatrix}5 & 8 & 4\\ 8 & 3 & -1\\ 4 & -1 & 2\end{bmatrix}\begin{bmatrix}1.00\\ -4.30\\ 8.10\end{bmatrix}=\begin{bmatrix}3.00\\ -13.00\\ 24.50\end{bmatrix}$$

如果 Ax 近似于 x 的一个倍数，则两个向量中对应元素的比值应该接近常数. 所以计算：

{Ax 中的元素} ÷	{x 中的元素} =	{比值}
3.00	1.00	3.000
−13.00	−4.30	3.023
24.50	8.10	3.025

Ax 中的每个元素大约为 x 中对应元素的3倍，所以 x 近似于一个本征向量. 上面比值中的任

何一个都是本征值的估计(保留五位小数,本征值为 3.024 09).

第 5 章补充题

在这些补充题中，$\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是具有适中维度的方阵.

1. 判断每个命题的真假，并说明理由.
   a. 如果 $\boldsymbol{A}$ 可逆且 1 是 $\boldsymbol{A}$ 的本征值，则 1 也是 $\boldsymbol{A}^{-1}$ 的本征值.
   b. 如果 $\boldsymbol{A}$ 行等价于单位矩阵 $\boldsymbol{I}$，则 $\boldsymbol{A}$ 可对角化.
   c. 如果 $\boldsymbol{A}$ 有一行或一列为零，则 0 为 $\boldsymbol{A}$ 的本征值.
   d. $\boldsymbol{A}$ 的每个本征值都是 $\boldsymbol{A}^2$ 的本征值.
   e. $\boldsymbol{A}$ 的每个本征向量都是 $\boldsymbol{A}^2$ 的本征向量.
   f. 可逆矩阵 $\boldsymbol{A}$ 的每个本征向量都是 $\boldsymbol{A}^{-1}$ 的本征向量.
   g. 本征值必须为非零数量.
   h. 本征向量必须是非零向量.
   i. 对应于同一个本征值的两个本征向量总是线性相关的.
   j. 相似矩阵的本征值相同.
   k. 相似矩阵的本征向量相同.
   l. 矩阵 $\boldsymbol{A}$ 的两个本征向量的和仍为 $\boldsymbol{A}$ 的本征向量.

370

   m. 上三角矩阵 $\boldsymbol{A}$ 的本征值为其对角线上的非零元素.
   n. 若计重数，矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{A}^{\mathrm{T}}$ 的本征值相同.
   o. 如果 5×5 矩阵 $\boldsymbol{A}$ 不同的本征值少于 5 个，则 $\boldsymbol{A}$ 不可对角化.
   p. 存在一个在 $\mathbf{R}^2$ 中没有本征向量的 2×2 矩阵.
   q. 如果 $\boldsymbol{A}$ 可对角化，则 $\boldsymbol{A}$ 的列向量线性无关.
   r. 一个非零向量不能对应于 $\boldsymbol{A}$ 的两个不同的本征值.
   s. 矩阵(方阵)$\boldsymbol{A}$ 可逆当且仅当存在一个坐标系，变换 $\boldsymbol{x}\to\boldsymbol{Ax}$ 在这个坐标系中可由一个对角矩阵表示.
   t. 如果 $\mathbf{R}^n$ 的标准基中每一个向量 $\boldsymbol{e}_j$ 都是 $\boldsymbol{A}$ 的本征向量，则 $\boldsymbol{A}$ 为对角矩阵.
   u. 如果 $\boldsymbol{A}$ 与一个可对角化的矩阵 $\boldsymbol{B}$ 相似，则 $\boldsymbol{A}$ 也可对角化.
   v. 如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为可逆的 $n\times n$ 矩阵，则 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 相似.
   w. 有 n 个线性无关本征向量的 $n\times n$ 矩阵 $\boldsymbol{A}$ 是可逆的.
   x. 如果 $\boldsymbol{A}$ 为可对角化的 $n\times n$ 矩阵，则 $\mathbf{R}^n$ 中每个向量都可以写成 $\boldsymbol{A}$ 的本征向量的线性组合.
2. 证明如果 $\boldsymbol{x}$ 为矩阵乘积 $\boldsymbol{AB}$ 的本征向量，且 $\boldsymbol{Bx}\neq\boldsymbol{0}$，则 $\boldsymbol{Bx}$ 为 $\boldsymbol{BA}$ 的本征向量.
3. 设 $\boldsymbol{x}$ 是矩阵 $\boldsymbol{A}$ 的本征向量，对应于本征值 λ.
   a. 证明 $\boldsymbol{x}$ 为 $5\boldsymbol{I}-\boldsymbol{A}$ 的本征向量. 求对应的本征值.
   b. 证明 $\boldsymbol{x}$ 为 $5\boldsymbol{I}-3\boldsymbol{A}+\boldsymbol{A}^2$ 的本征向量. 求对应的本征值.
4. 用数学归纳法证明：如果 λ 是 $n\times n$ 矩阵 $\boldsymbol{A}$ 的本征值，对应的本征向量是 $\boldsymbol{x}$，则对于每个正整数 m，λ^m 是 $\boldsymbol{A}^m$ 的本征值，对应的本征向量是 $\boldsymbol{x}$.
5. 如果 $p(t)=c_0+c_1t+c_2t^2+\cdots+c_nt^n$，用 $\boldsymbol{A}$ 的幂代替 $p(t)$ 中相应的 t 的幂得到的矩阵定义为 $p(\boldsymbol{A})(\boldsymbol{A}^0=\boldsymbol{I})$. 即
$$p(\boldsymbol{A})=c_0\boldsymbol{I}+c_1\boldsymbol{A}+c_2\boldsymbol{A}^2+\cdots+c_n\boldsymbol{A}^n$$
   证明如果 λ 为 $\boldsymbol{A}$ 的本征值，则 $p(\lambda)$ 为 $p(\boldsymbol{A})$ 的本征值.
6. 设 $\boldsymbol{A}=\boldsymbol{PDP}^{-1}$，其中 $\boldsymbol{P}$ 为 2×2 矩阵且 $\boldsymbol{D}=\begin{bmatrix}2&0\\0&7\end{bmatrix}$.

a. 令 $\boldsymbol{B}=5\boldsymbol{I}-3\boldsymbol{A}+\boldsymbol{A}^2$. 通过求 $\boldsymbol{B}$ 的适当分解，证明 $\boldsymbol{B}$ 可对角化.

b. 已知 $p(t)$ 和 $p(\boldsymbol{A})$ 如补充题5中所示，证明 $p(\boldsymbol{A})$ 可对角化.

7. 设 $\boldsymbol{A}$ 可对角化且 $p(t)$ 为 $\boldsymbol{A}$ 的特征方程. $p(\boldsymbol{A})$ 的定义如补充题5中所述，证明 $p(\boldsymbol{A})$ 为零矩阵. 这个事实对任何方阵都是正确的，称为凯莱-哈密顿定理.

8. a. 设 $\boldsymbol{A}$ 为可对角化的 $n\times n$ 矩阵. 证明如果本征值 λ 的重数是 n，则 $\boldsymbol{A}=\lambda\boldsymbol{I}$.

b. 根据(a)证明矩阵 $\boldsymbol{A}=\begin{bmatrix}3&1\\0&3\end{bmatrix}$ 不可对角化.

9. 证明当 $\boldsymbol{A}$ 所有的本征值的绝对值都小于1时，$\boldsymbol{I}-\boldsymbol{A}$ 可逆.

10. 证明如果 $\boldsymbol{A}$ 可对角化，且所有的本征值长度都小于1，则当 $k\to\infty$ 时，$\boldsymbol{A}^k$ 趋于零矩阵. [提示：考虑 $\boldsymbol{A}^k\boldsymbol{x}$，其中 $\boldsymbol{x}$ 表示 $\boldsymbol{I}$ 的任一列向量.]

11. 设 $\boldsymbol{u}$ 为 $\boldsymbol{A}$ 的本征向量，对应于本征值 λ，并令 H 为 $\mathbf{R}^n$ 中通过 $\boldsymbol{u}$ 和原点的直线.

a. 解释 H 在 $\boldsymbol{A}$ 下不变的原因，即只要 $\boldsymbol{x}$ 属于 H，$\boldsymbol{Ax}$ 就属于 H.

b. 令 K 为 $\mathbf{R}^n$ 中的一维子空间，且在 $\boldsymbol{A}$ 下不变. 解释 K 包含 $\boldsymbol{A}$ 的一个本征向量的原因.

12. 设 $\boldsymbol{G}=\begin{bmatrix}\boldsymbol{A}&\boldsymbol{X}\\0&\boldsymbol{B}\end{bmatrix}$. 利用5.2节中行列式公式(1)，解释 $\det\boldsymbol{G}=(\det\boldsymbol{A})(\det\boldsymbol{B})$ 的原因. 并根据这个公式证明 $\boldsymbol{G}$ 的特征多项式是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的特征多项式的乘积.

利用补充题12求下列矩阵的本征值：

13. $\boldsymbol{A}=\begin{bmatrix}3&-2&8\\0&5&-2\\0&-4&3\end{bmatrix}$　　14. $\boldsymbol{A}=\begin{bmatrix}1&5&-6&-7\\2&4&5&2\\0&0&-7&-4\\0&0&3&1\end{bmatrix}$

371

15. 令 $\boldsymbol{J}$ 为所有元素都是1的 $n\times n$ 矩阵，并考虑 $\boldsymbol{A}=(a-b)\boldsymbol{I}+b\boldsymbol{J}$；即

$$\boldsymbol{A}=\begin{bmatrix}a&b&b&\cdots&b\\b&a&b&\cdots&b\\b&b&a&\cdots&b\\\vdots&\vdots&\vdots&\ddots&\vdots\\b&b&b&\cdots&a\end{bmatrix}$$

利用第三章补充题16中的结论证明 $\boldsymbol{A}$ 的本征值是 $a-b$ 和 $a+(n-1)b$. 它们各自的重数是多少？

16. 利用补充题15的结果求矩阵 $\begin{bmatrix}1&2&2\\2&1&2\\2&2&1\end{bmatrix}$ 和 $\begin{bmatrix}7&3&3&3&3\\3&7&3&3&3\\3&3&7&3&3\\3&3&3&7&3\\3&3&3&3&7\end{bmatrix}$ 的本征值.

17. 设 $\boldsymbol{A}=\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}$. 回忆5.4节习题25中，$\operatorname{tr}\boldsymbol{A}$（$\boldsymbol{A}$ 的迹）是 $\boldsymbol{A}$ 中对角线元素的和. 证明 $\boldsymbol{A}$ 的特征多项式是

$$\lambda^2-(\operatorname{tr}\boldsymbol{A})\lambda+\det\boldsymbol{A}$$

并证明 2×2 矩阵 $\boldsymbol{A}$ 的本征值都是实数当且仅当 $\det\boldsymbol{A}\leqslant\left(\dfrac{\operatorname{tr}\boldsymbol{A}}{2}\right)^2$.

18. 设 $A=\begin{bmatrix}0.4 & -0.3\\ 0.4 & 1.2\end{bmatrix}$. 解释当 $k\to\infty$ 时，A^k 趋于 $\begin{bmatrix}-0.5 & -0.75\\ 1.0 & 1.50\end{bmatrix}$.

补充题 19 ~ 23 考虑的是多项式 $p(t)=a_0+a_1t+\cdots+a_{n-1}t^{n-1}+t^n$ 及称为 p 的**伴随矩阵**(companion matrix)的 $n\times n$ 矩阵 C_p:

$$C_p=\begin{bmatrix}0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & & 0\\ \vdots & & & & \vdots\\ 0 & 0 & 0 & & 1\\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1}\end{bmatrix}$$

19. 写出 $p(t)=6-5t+t^2$ 的伴随矩阵 C_p,然后求 C_p的特征多项式.
20. 设
$$p(t)=(t-2)(t-3)(t-4)=-24+26t-9t^2+t^3$$
写出 $p(t)$的伴随矩阵，并利用第 3 章中的技巧求解其特征多项式.
21. 用数学归纳法证明，对于 $n\geqslant 2$,
$$\det(C_p-\lambda I)=(-1)^n(a_0+a_1\lambda+\cdots+a_{n-1}\lambda^{n-1}+\lambda^n)=(-1)^n p(\lambda)$$
[提示:按第一列展开,证明 $\det(C_p-\lambda I)$形如$(-\lambda)B+(-1)^n a_0$,其中 B 为某个特定的多项式(由归纳假设).]
22. 设 $p(t)=a_0+a_1t+a_2t^2+t^3$,$p(\lambda)$等于零.
   a. 写出 p 的伴随矩阵.
   b. 解释 $\lambda^3=-a_0-a_1\lambda-a_2\lambda^2$ 的原因，并证明$(1,\lambda,\lambda^2)$为 p 的伴随矩阵的本征向量.
23. 设 p 为补充题 22 中的多项式，且方程 $p(t)=0$ 有不同的根 λ_1，λ_2，λ_3. 令 V 为范德蒙德矩阵
$$V=\begin{bmatrix}1 & 1 & 1\\ \lambda_1 & \lambda_2 & \lambda_3\\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2\end{bmatrix}$$
(V 的转置在第 2 章补充题 11 中讨论过.)利用补充题 22 和本章的一个定理证明 V 可逆(但并不计算 V^{-1}). 然后解释 $V^{-1}C_pV$ 为对角矩阵的原因.
24. MATLAB 中的命令 `roots(p)`可以计算多项式方程 $p(t)=0$ 的根. 阅读一本 MATLAB 手册，然后描述 `roots` 命令所用算法的基本思想.
25. [M]如果可能，利用矩阵软件对下面的矩阵进行对角化.
$$A=\begin{bmatrix}-3 & -2 & 0\\ 14 & 7 & -1\\ -6 & -3 & 1\end{bmatrix}$$
用本征值命令生成对角矩阵 D. 如果软件中有可以直接生成本征向量的命令,用它来生成可逆矩阵 P. 然后计算 $AP-PD$ 和 PDP^{-1}. 讨论你的结果.
26. 对 $A=\begin{bmatrix}-8 & 5 & -2 & 0\\ -5 & 2 & 1 & -2\\ 10 & -8 & 6 & -3\\ 3 & -2 & 1 & 0\end{bmatrix}$重做补充题 25.

372

第6章　正交性与最小二乘法

实例介绍：北美大地基准的校正

试想这样一项工程，它预计花费十年时间、耗费许多人力来建立并求解一个拥有900 000个未知量、1 800 000个方程的方程组. 1974年，当美国大地测量调查局着手校正北美大地基准(NAD)时，情况正是如此. 他们需要绘制出一张布满268 000个经仔细测量并且标记的点的网络，它覆盖整个北美洲，跨越巴拿马地峡、格陵兰岛、夏威夷群岛、维京群岛、波多黎各群岛和加勒比海群岛.

所有土地测量、地图绘制、合法地产界的划定、州和地方土地利用的规划以及城市工程建设项目(如高速公路、公用交通干线)的规划都以大地基准中记录的经纬度为依据，这些数据只容许有厘米级的误差. 然而，自1927年上一次大地基准参考点校正以后，有200 000个以上新的点被增补到原有体系中. 况且随着时间推移，旧数据的误差越来越大，某些地区的地表运动每年多达5厘米. 到1970年，彻底更新大地基准的需求变得十分迫切，确定参考点新坐标的计划随即出台.

大地基准校正过程中，收集到的测量数据时间跨度长达140年，这些数据都需要被标准化并且转换成计算机可以读取的格式(例如,计算机读取加利福尼亚州圣安地列斯断层一带几年前的测量数据后,再利用地壳运动的数学模型,就能得到这一区域现在的数据). 接下来，测量数据还需要反复核对，以排除原始数据本身的错误或是在录入计算机时出的错. 最后的计算过程涉及约1 800 000个观测量，根据其相对精度每一个量都被赋予权重，并且每一个量都定义了一个方程.

求解北美大地基准的方程组在通常意义下无解，但有*最小二乘解*，从某种意义上说，它赋给参考点的经纬度值与这1 800 000个观测量最为吻合. 这个最小二乘解是通过求解一个相关线性方程组而得到，其中包含了928 735个正规方程以及928 735个变量.

373

这些正规方程对现有计算机来说仍然十分庞大，因此需要把原方程组分解成许多小的方程组. 分解的方法是递归应用*赫尔墨特分块法*，将方程的系数矩阵逐步划分成越来越小的若干块. 最小的矩阵分块给出的方程覆盖了含500~2000个大地基准参考点的连续地理区域. 图6-1描述了利用赫尔墨特分块法将美国本土细分的过程. 求得小方程组的解后，经过若干中间步骤最终可以得到全部928 735个变量的值. [1]

---

1. 赫尔墨特分块策略的数学讨论，以及整个北美大地基准项目的细节都收录于 *North American Datum of* 1983, Charles R. Schwarz (ed.), National Geodetic Survey, National Oceanic and Atmospheric Administration(NOAA) Professional Paper NOS 2, 1989.

北美大地基准校正数据库完成于 1983 年. 三年后，经过广泛分析和超过 940 小时的计算机运行，这个当时遇到的规模最大的最小二乘问题终于得到了解决.

如以上实例所述，由实验数据得到的线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 经常没有解. 这时可取的替代方法往往是找一个向量 $\hat{\boldsymbol{x}}$，使 $\boldsymbol{A}\hat{\boldsymbol{x}}$ 与 $\boldsymbol{b}$ 的距离尽可能最小. 这里的距离是一个平方和，6.1 节中将给出其具体定义. 满足前述条件的向量 $\hat{\boldsymbol{x}}$ 称作 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小二乘解. 本章中，6.1 至 6.3 节给出正交性和正交投影的基本概念，6.5 节则运用这些概念求最小二乘解 $\hat{\boldsymbol{x}}$.

374

图 6-1 赫尔墨特分块后的美国国土

矩阵分解是数值线性代数中广泛使用的技巧. 6.4 节通过建立矩阵分解为我们理解正交投影提供了另一种视角. 本章其余各节考察了若干源于应用的最小二乘问题，有的问题讨论范围是比 $\mathbf{R}^n$ 更一般的向量空间. 但不论何种情形，这些向量空间中的数量都是实数.

6.1 内积、长度和正交性

长度、距离和垂直这些在 $\mathbf{R}^2$，$\mathbf{R}^3$ 中熟知的几何概念在本节中被推广到 $\mathbf{R}^n$ 上. 这些概念为解决许多实际问题，包括前述的最小二乘问题，提供了强大的几何工具. 以下我们采用向量内积的形式给出它们的定义.

6.1.1 内积

设 $\boldsymbol{u}$，$\boldsymbol{v}$ 是 $\mathbf{R}^n$ 中的向量，我们将 $\boldsymbol{u}$，$\boldsymbol{v}$ 看作 $n\times 1$ 矩阵. 则转置矩阵 $\boldsymbol{u}^{\mathrm{T}}$ 是 $1\times n$ 矩阵，矩阵乘积 $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}$ 是 1×1 矩阵，后者我们常记为不带方括号的某个实数(数量). 我们称数值 $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}$ 为 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的**内积**(inner product)，常记作 $\boldsymbol{u}\cdot\boldsymbol{v}$. 在 2.1 节习题中提到过的这个内积，有时也称作**点积**(dot product). 如果

$$\boldsymbol{u}=\begin{bmatrix}u_1\\u_2\\\vdots\\u_n\end{bmatrix} \quad 且 \quad \boldsymbol{v}=\begin{bmatrix}v_1\\v_2\\\vdots\\v_n\end{bmatrix}$$

则 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的内积是

$$\begin{bmatrix}u_1 & u_2\cdots u_n\end{bmatrix}\begin{bmatrix}v_1\\v_2\\\vdots\\v_n\end{bmatrix}=u_1v_1+u_2v_2+\cdots+u_nv_n$$

375 **【例题1】** $\boldsymbol{u}=\begin{bmatrix}2\\-5\\-1\end{bmatrix}$, $\boldsymbol{v}=\begin{bmatrix}3\\2\\-3\end{bmatrix}$, 计算 $\boldsymbol{u}\cdot\boldsymbol{v}$ 和 $\boldsymbol{v}\cdot\boldsymbol{u}$ 的值.

解:

$$\boldsymbol{u}\cdot\boldsymbol{v}=\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}=[2\quad -5\quad -1]\begin{bmatrix}3\\2\\-3\end{bmatrix}=(2)(3)+(-5)(2)+(-1)(-3)=-1$$

$$\boldsymbol{v}\cdot\boldsymbol{u}=\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u}=[3\quad 2\quad -3]\begin{bmatrix}2\\-5\\-1\end{bmatrix}=(3)(2)+(2)(-5)+(-3)(-1)=-1$$ ■

例题1的计算过程清楚地说明了 $\boldsymbol{u}\cdot\boldsymbol{v}$ 和 $\boldsymbol{v}\cdot\boldsymbol{u}$ 相等的原因. 一般来说，内积的交换性都成立. 下列内积性质易由2.1节矩阵转置运算的性质推得(见本节习题21和22).

【定理1】 设 $\boldsymbol{u}$, $\boldsymbol{v}$ 和 $\boldsymbol{w}$ 都是 $\mathbf{R}^n$ 中的向量，c 是数量. 则:

a. $\boldsymbol{u}\cdot\boldsymbol{v}=\boldsymbol{v}\cdot\boldsymbol{u}$

b. $(\boldsymbol{u}+\boldsymbol{v})\cdot\boldsymbol{w}=\boldsymbol{u}\cdot\boldsymbol{w}+\boldsymbol{v}\cdot\boldsymbol{w}$

c. $(c\boldsymbol{u})\cdot\boldsymbol{v}=c(\boldsymbol{u}\cdot\boldsymbol{v})=\boldsymbol{u}\cdot(c\boldsymbol{v})$

d. $\boldsymbol{u}\cdot\boldsymbol{u}\geqslant 0$，且 $\boldsymbol{u}\cdot\boldsymbol{u}=0$ 当且仅当 $\boldsymbol{u}=\mathbf{0}$

将定理1中的性质(b)和(c)结合运用多次，可得如下实用法则:

$$(c_1\boldsymbol{u}_1+\cdots+c_p\boldsymbol{u}_p)\cdot\boldsymbol{w}=c_1(\boldsymbol{u}_1\cdot\boldsymbol{w})+\cdots+c_p(\boldsymbol{u}_p\cdot\boldsymbol{w})$$

6.1.2　向量长度

设 $\boldsymbol{v}$ 是 $\mathbf{R}^n$ 中的向量，其中元素为 $v_1,\cdots,v_n$. 由于 $\boldsymbol{v}\cdot\boldsymbol{v}$ 非负，我们可以定义 $\boldsymbol{v}\cdot\boldsymbol{v}$ 的平方根.

【定义】 向量 $\boldsymbol{v}$ 的**长度**(length)或**范式**(norm)是非负数量 $\|\boldsymbol{v}\|$，其定义为:

$$\|\boldsymbol{v}\|=\sqrt{\boldsymbol{v}\cdot\boldsymbol{v}}=\sqrt{v_1^2+v_2^2+\cdots+v_n^2}\text{，且有}\|\boldsymbol{v}\|^2=\boldsymbol{v}\cdot\boldsymbol{v}$$

若 $\boldsymbol{v}$ 是 $\mathbf{R}^2$ 中的向量，设 $\boldsymbol{v}=\begin{bmatrix}a\\b\end{bmatrix}$. 如果我们把 $\boldsymbol{v}$ 看成是平面上的几何点，则根据勾股定理，$\|\boldsymbol{v}\|$ 恰好是原点到 $\boldsymbol{v}$ 的线段长度，如图6-2所示.

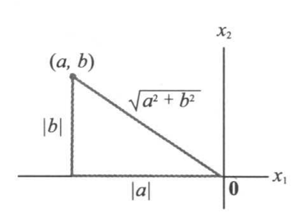


图6-2　将 $\|\boldsymbol{v}\|$ 解释为长度

376 类似地，通过计算立方体对角线可知，$\mathbf{R}^3$ 中向量 $\boldsymbol{v}$ 的长度定义与通常的长度概念一致.

对任意数量 c，$c\boldsymbol{v}$ 的长度是 $\boldsymbol{v}$ 长度的 $|c|$ 倍，即

$$\|c\boldsymbol{v}\|=|c|\|\boldsymbol{v}\|$$

(为证明这个等式，只需计算 $\|c\boldsymbol{v}\|^2=(c\boldsymbol{v})\cdot(c\boldsymbol{v})=c^2\boldsymbol{v}\cdot\boldsymbol{v}=|c|^2\|\boldsymbol{v}\|^2$，再取平方根即可).

长度为1的向量称作**单位向量**(unit vector). 如果让非零向量 $\boldsymbol{v}$ 除以它的长度(即乘以 $1/\|\boldsymbol{v}\|$)，我们得到单位向量 $\boldsymbol{u}$，因为 $\boldsymbol{u}$ 的长度等于 $(1/\|\boldsymbol{v}\|)\|\boldsymbol{v}\|$. 这种由 $\boldsymbol{v}$ 构造 $\boldsymbol{u}$ 的过程有时称作 $\boldsymbol{v}$ 的**正规化**(normalizing)，同时称 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 同向.

为节省篇幅，下面的例子中采用了(列)向量的简单记法.

【例题2】 令 $\boldsymbol{v}=(1,-2,2,0)$，找出与 $\boldsymbol{v}$ 同向的单位向量 $\boldsymbol{u}$.

解： 首先计算 $\boldsymbol{v}$ 的长度：

$$\|\boldsymbol{v}\|^2=\boldsymbol{v}\cdot\boldsymbol{v}=(1)^2+(-2)^2+(2)^2+(0)^2=9$$

$$\|\boldsymbol{v}\|=\sqrt{9}=3$$

然后用 $1/\|\boldsymbol{v}\|$ 乘 $\boldsymbol{v}$，得到：

$$\boldsymbol{u}=\frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}=\frac{1}{3}\boldsymbol{v}=\frac{1}{3}\begin{bmatrix}1\\-2\\2\\0\end{bmatrix}=\begin{bmatrix}1/3\\-2/3\\2/3\\0\end{bmatrix}$$

为验证 $\|\boldsymbol{u}\|=1$，只需证明 $\|\boldsymbol{u}\|^2=1$.

$$\|\boldsymbol{u}\|^2=\boldsymbol{u}\cdot\boldsymbol{u}=\left(\frac{1}{3}\right)^2+\left(-\frac{2}{3}\right)^2+\left(\frac{2}{3}\right)^2+(0)^2=\frac{1}{9}+\frac{4}{9}+\frac{4}{9}+0=1$$ ■

【例题3】 设 W 是 $\mathbf{R}^2$ 中由 $\boldsymbol{x}=(2/3,1)$ 张成的子空间. 找一单位向量 $\boldsymbol{z}$，要求 $\boldsymbol{z}$ 是 W 的基.

解： 如图 6-3a 所示，W 由 $\boldsymbol{x}$ 与全体数量的乘积组成. W 中的任意非零向量都是 W 的基. 为了简化计算过程，适当缩放 $\boldsymbol{x}$ 以消去其中的分数，即，用 3 乘 $\boldsymbol{x}$ 得到

$$\boldsymbol{y}=\begin{bmatrix}2\\3\end{bmatrix}$$

现在我们计算 $\|\boldsymbol{y}\|^2=2^2+3^2=13$，$\|\boldsymbol{y}\|^2=\sqrt{13}$，并且正规化向量 $\boldsymbol{y}$ 可得

$$\boldsymbol{z}=\frac{1}{\sqrt{13}}\begin{bmatrix}2\\3\end{bmatrix}=\begin{bmatrix}2/\sqrt{13}\\3/\sqrt{13}\end{bmatrix}$$

如图 6-3b 所示. 另一个满足条件的单位向量是 $(-2/\sqrt{13},-3/\sqrt{13})$. ■ 377

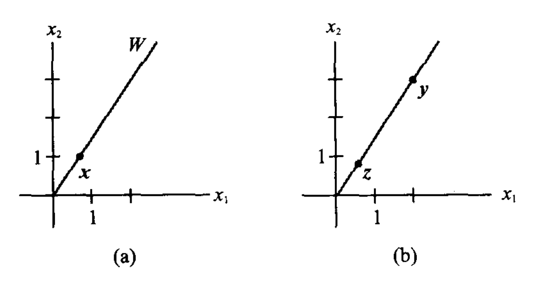


图 6-3 将向量正规化以得到单位向量

6.1.3 $\mathbf{R}^n$ 中的距离

现在我们可以描述两个向量靠近的程度. 回顾一下，如果 a 和 b 是实数，则它们在数轴上的距离就是数值 $|a-b|$. 图 6-4 给出了两个具体的例子. $\mathbf{R}$ 中这种距离的定义在 $\mathbf{R}^n$ 中也有自然的推广.

【定义】 设 $\boldsymbol{u}$，$\boldsymbol{v}$ 是 $\mathbf{R}^n$ 中的向量．**$\boldsymbol{u}$ 和 $\boldsymbol{v}$ 之间的距离**(distance between $\boldsymbol{u}$ and $\boldsymbol{v}$)就是向量 $\boldsymbol{u}-\boldsymbol{v}$ 的长度，记作 $\text{dist}(\boldsymbol{u},\boldsymbol{v})$．即

$$\text{dist}(\boldsymbol{u},\boldsymbol{v}) = \|\boldsymbol{u}-\boldsymbol{v}\|$$

在 $\mathbf{R}^2$ 和 $\mathbf{R}^3$ 中，距离的上述定义和两点之间常用的欧几里德距离公式是一致的，接下来的两个例子说明了这一点．

【例题4】 计算向量 $\boldsymbol{u}=(7,1)$ 和 $\boldsymbol{v}=(3,2)$ 之间的距离．

解： 计算

$$\boldsymbol{u}-\boldsymbol{v} = \begin{bmatrix}7\\1\end{bmatrix} - \begin{bmatrix}3\\2\end{bmatrix} = \begin{bmatrix}4\\-1\end{bmatrix}$$

$$\|\boldsymbol{u}-\boldsymbol{v}\| = \sqrt{4^2+(-1)^2} = \sqrt{17}$$

向量 $\boldsymbol{u}$，$\boldsymbol{v}$ 和 $\boldsymbol{u}-\boldsymbol{v}$ 如图6-5所示，$\boldsymbol{v}$ 加上 $\boldsymbol{u}-\boldsymbol{v}$ 得 $\boldsymbol{u}$．注意，图6-5中的平行四边形表明 $\boldsymbol{u}$ 到 $\boldsymbol{v}$ 的距离等于 $\boldsymbol{u}-\boldsymbol{v}$ 到 $\boldsymbol{0}$ 的距离．■

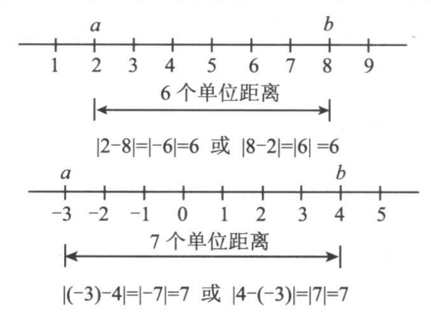


图6-4　**R** 中的距离

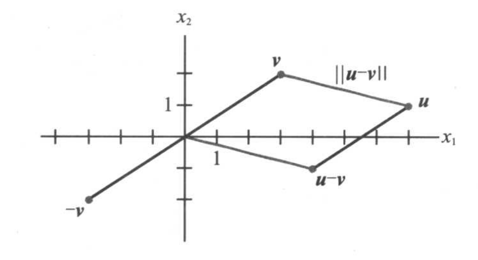


图6-5　$\boldsymbol{u}$ 和 $\boldsymbol{v}$ 之间的距离就是 $\boldsymbol{u}-\boldsymbol{v}$ 的长度

378

【例题5】 若 $\boldsymbol{u}=(u_1,u_2,u_3)$ 且 $\boldsymbol{v}=(v_1,v_2,v_3)$，则

$$\begin{aligned}\text{dist}(\boldsymbol{u},\boldsymbol{v}) &= \|\boldsymbol{u}-\boldsymbol{v}\| = \sqrt{(\boldsymbol{u}-\boldsymbol{v})\cdot(\boldsymbol{u}-\boldsymbol{v})} \\ &= \sqrt{(u_1-v_1)^2+(u_2-v_2)^2+(u_3-v_3)^2}\end{aligned}$$

■

6.1.4　正交向量

本章余下的内容都依赖一个事实，这就是：欧氏几何中垂线的概念可以推广到 $\mathbf{R}^n$ 空间中．

考虑 $\mathbf{R}^2$ 或 $\mathbf{R}^3$ 空间以及由向量 $\boldsymbol{u}$，$\boldsymbol{v}$ 确定的过原点的两条直线．如图6-6，这两条直线互相垂直当且仅当 $\boldsymbol{u}$ 到 $\boldsymbol{v}$ 的距离等于 $\boldsymbol{u}$ 到 $-\boldsymbol{v}$ 的距离，当且仅当这两个距离的平方相等．于是，

$$\begin{aligned}[\text{dist}(\boldsymbol{u},-\boldsymbol{v})]^2 &= \|\boldsymbol{u}-(-\boldsymbol{v})\|^2 = \|\boldsymbol{u}+\boldsymbol{v}\|^2 \\ &= (\boldsymbol{u}+\boldsymbol{v})\cdot(\boldsymbol{u}+\boldsymbol{v}) \\ &= \boldsymbol{u}\cdot(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{v}\cdot(\boldsymbol{u}+\boldsymbol{v}) && \text{定理1(b)} \\ &= \boldsymbol{u}\cdot\boldsymbol{u}+\boldsymbol{u}\cdot\boldsymbol{v}+\boldsymbol{v}\cdot\boldsymbol{u}+\boldsymbol{v}\cdot\boldsymbol{v} && \text{定理1(a),(b)} \\ &= \|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2+2\boldsymbol{u}\cdot\boldsymbol{v} && \text{定理1(a)}\end{aligned} \tag{1}$$

将 $\boldsymbol{v}$ 替换成 $-\boldsymbol{v}$，同样的推导表明

$$[\text{dist}(\boldsymbol{u},\boldsymbol{v})]^2 = \|\boldsymbol{u}\|^2 + \|-\boldsymbol{v}\|^2 + 2\boldsymbol{u}\cdot(-\boldsymbol{v}) = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - 2\boldsymbol{u}\cdot\boldsymbol{v}$$

所以，这两个距离的平方相等当且仅当 $2\boldsymbol{u}\cdot\boldsymbol{v} = -2\boldsymbol{u}\cdot\boldsymbol{v}$，当且仅当 $\boldsymbol{u}\cdot\boldsymbol{v}=0$.

上述计算说明，当把 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 看成几何点时，分别通过 $\boldsymbol{u}$，$\boldsymbol{v}$ 及原点的两条直线互相垂直当且仅当 $\boldsymbol{u}\cdot\boldsymbol{v}=0$. 下面的定义把垂直（或称正交，后者在线性代数中使用更为普遍）的概念推广到 $\mathbf{R}^n$.

【定义】 称 $\mathbf{R}^n$ 中两个向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$（互相）**正交**（orthogonal），如果 $\boldsymbol{u}\cdot\boldsymbol{v}=0$.

注意零向量正交于 $\mathbf{R}^n$ 中任意向量，因为对任意 $\boldsymbol{v}$ 都有 $\mathbf{0}^{\mathrm{T}}\boldsymbol{v}=0$.

下面的定理给出了关于正交向量的一个重要事实，其证明可以利用前文公式(1)的推导过程以及正交的定义来完成. 图 6-7 右侧的三角形标记出了这个定理涉及的三个长度.

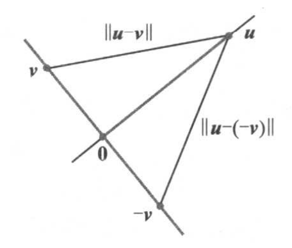


图 6-6 $\boldsymbol{u}$ 与 $\boldsymbol{v}$ 互相垂直当且仅当 $\|\boldsymbol{u}-\boldsymbol{v}\| = \|\boldsymbol{u}-(-\boldsymbol{v})\|$

图 6-7 与勾股定理有关的三个长度

379

【定理 2】 勾股定理

两向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 正交当且仅当 $\|\boldsymbol{u}+\boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2$.

6.1.5 正交补

作为内积的另一个应用，这里我们引入一个新的概念，它将在 6.3 节和本章其他部分用到. 如果向量 $\boldsymbol{z}$ 正交于 $\mathbf{R}^n$ 的子空间 W 中的任意向量，则称 $\boldsymbol{z}$ **正交于**（orthogonal to）W. 所有正交于 W 的向量 $\boldsymbol{z}$ 构成的集合称作 W **的正交补**（orthogonal complement），记作 $W^{\perp}$（读作 W 正交）.

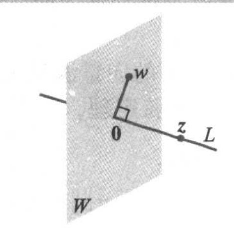


图 6-8 过原点的互为正交补的平面和直线

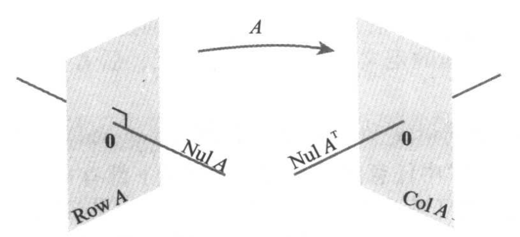
【例题 6】 令 W 是 $\mathbf{R}^3$ 中过原点的一个平面，L 是过原点且与 W 正交的直线. 如图6-8，设 $\boldsymbol{z}$，$\boldsymbol{w}$ 是分别属于 L 和 W 的非零向量，则从 $\mathbf{0}$ 到 $\boldsymbol{z}$ 的线段与从 $\mathbf{0}$ 到 $\boldsymbol{w}$ 的线段垂直，即 $\boldsymbol{z}\cdot\boldsymbol{w}=0$. 因此 L 上的任意向量都正交于 W 中的任意向量 $\boldsymbol{w}$. 事实上，L 包含了所有与 W 中全体 $\boldsymbol{w}$ 正交的向量，且 W 包含了所有与 L 中全体 $\boldsymbol{z}$ 正交的向量，即，

$$L = W^{\perp},\ W = L^{\perp}$$ ■

设 W 是 $\mathbf{R}^n$ 中子空间，下面关于 $W^{\perp}$ 的两个事实将在本章后继内容中用到，其证明可参考习题 29 和 30. 习题 27 ~ 31 都是关于内积性质的很好的练习.

1. 向量 $\boldsymbol{x}$ 属于 $W^{\perp}$ 当且仅当 $\boldsymbol{x}$ 正交于张成 W 的某个集合中的每一向量.
2. $W^{\perp}$ 是 $\mathbf{R}^n$ 的子空间.

下面的定理以及习题31证实了4.6节中有关图6-9所示子空间的断言(见4.6节习题28).



380

图6-9　$m\times n$矩阵A定义的基本空间

【定理3】　令$\boldsymbol{A}$是一个$m\times n$矩阵，则$\boldsymbol{A}$的行向量空间的正交补是$\boldsymbol{A}$的解空间，$\boldsymbol{A}$的列向量空间的正交补是$\boldsymbol{A}^{\mathrm{T}}$的解空间，即，

$$(\mathrm{Row}\,\boldsymbol{A})^{\perp}=\mathrm{Nul}\,\boldsymbol{A},\ (\mathrm{Col}\,\boldsymbol{A})^{\perp}=\mathrm{Nul}\,\boldsymbol{A}^{\mathrm{T}}$$

证明：由$\boldsymbol{Ax}$的行列计算法则可知，若$\boldsymbol{x}\in\mathrm{Nul}\,\boldsymbol{A}$，则$\boldsymbol{x}$与$\boldsymbol{A}$的每一行(看作$\mathbf{R}^n$中的向量)都正交. 由于$\boldsymbol{A}$的行向量空间是由这些行向量张成的，因此$\boldsymbol{x}$与Row $\boldsymbol{A}$正交. 反过来，如果$\boldsymbol{x}$与Row $\boldsymbol{A}$正交，$\boldsymbol{x}$自然与$\boldsymbol{A}$的每一行向量正交，于是$\boldsymbol{Ax}=\mathbf{0}$. 这就完成了定理第一个结论的证明. 这一结论对任意矩阵都成立，亦对$\boldsymbol{A}^{\mathrm{T}}$成立，即$\boldsymbol{A}^{\mathrm{T}}$的行向量空间的正交补是$\boldsymbol{A}^{\mathrm{T}}$的解空间. 又因为Row $\boldsymbol{A}^{\mathrm{T}}$ = Col $\boldsymbol{A}$，所以第二个结论得证. ■

6.1.6　$\mathbf{R}^2$，$\mathbf{R}^3$中的角度(可选)

设$\boldsymbol{u}$，$\boldsymbol{v}$是$\mathbf{R}^2$或$\mathbf{R}^3$中的向量，考虑从原点出发分别经过$\boldsymbol{u}$，$\boldsymbol{v}$的两条直线所形成的夹角ϑ. 则ϑ与$\boldsymbol{u}$和$\boldsymbol{v}$的内积之间有着巧妙的联系，关系式为：

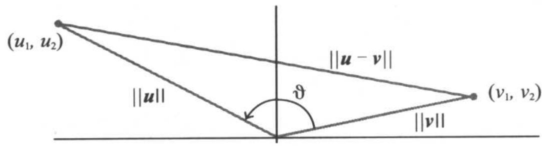
$$\boldsymbol{u}\cdot\boldsymbol{v}=\|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\vartheta \tag{2}$$

为验证$\mathbf{R}^2$中的向量满足上述公式，考察图6-10中以$\|\boldsymbol{u}\|$，$\|\boldsymbol{v}\|$和$\|\boldsymbol{u}-\boldsymbol{v}\|$为边的三角形. 根据余弦定理，

$$\|\boldsymbol{u}-\boldsymbol{v}\|^2=\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2-2\|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\vartheta$$

移项后可得：

$$\begin{aligned}\|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\vartheta&=\frac{1}{2}[\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2-\|\boldsymbol{u}-\boldsymbol{v}\|^2]\\&=\frac{1}{2}[u_1^2+u_2^2+v_1^2+v_2^2-(u_1-v_1)^2-(u_2-v_2)^2]\\&=u_1v_1+u_2v_2\\&=\boldsymbol{u}\cdot\boldsymbol{v}\end{aligned}$$



381

图6-10　向量之间的夹角

对$\mathbf{R}^3$的验证是类似的. 当$n>3$时，公式(2)同样可用来定义$\mathbf{R}^n$中两向量的夹角. 例

如在统计学中，对于适当的向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$，由公式(2)所定义的 $\cos\vartheta$ 的值被统计学家称作相关系数.

基础练习

$$\text{令 } \boldsymbol{a}=\begin{bmatrix}-2\\1\end{bmatrix},\ \boldsymbol{b}=\begin{bmatrix}-3\\1\end{bmatrix},\ \boldsymbol{c}=\begin{bmatrix}4/3\\-1\\2/3\end{bmatrix}\text{以及 } \boldsymbol{d}=\begin{bmatrix}5\\6\\-1\end{bmatrix}.$$

1. 计算 $\dfrac{\boldsymbol{a}\cdot\boldsymbol{b}}{\boldsymbol{a}\cdot\boldsymbol{a}}$ 和 $\left(\dfrac{\boldsymbol{a}\cdot\boldsymbol{b}}{\boldsymbol{a}\cdot\boldsymbol{a}}\right)\boldsymbol{a}$.
2. 求与 $\boldsymbol{c}$ 同向的单位向量 $\boldsymbol{u}$.
3. 证明 $\boldsymbol{d}$ 正交于 $\boldsymbol{c}$.
4. 利用基础练习 2 和 3 的结论解释为什么 $\boldsymbol{d}$ 一定正交于单位向量 $\boldsymbol{u}$.

习题 6.1

用下列向量完成习题 1 ~ 8 中各个量的计算：

$$\boldsymbol{u}=\begin{bmatrix}-1\\2\end{bmatrix},\ \boldsymbol{v}=\begin{bmatrix}4\\6\end{bmatrix},\ \boldsymbol{w}=\begin{bmatrix}3\\-1\\-5\end{bmatrix},\ \boldsymbol{x}=\begin{bmatrix}6\\-2\\3\end{bmatrix}$$

1. $\boldsymbol{u}\cdot\boldsymbol{u}$，$\boldsymbol{v}\cdot\boldsymbol{u}$ 及 $\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}}$　　2. $\boldsymbol{w}\cdot\boldsymbol{w}$，$\boldsymbol{x}\cdot\boldsymbol{w}$ 及 $\dfrac{\boldsymbol{x}\cdot\boldsymbol{w}}{\boldsymbol{w}\cdot\boldsymbol{w}}$　　3. $\dfrac{1}{\boldsymbol{w}\cdot\boldsymbol{w}}\boldsymbol{w}$　　4. $\dfrac{1}{\boldsymbol{u}\cdot\boldsymbol{u}}\boldsymbol{u}$

5. $\left(\dfrac{\boldsymbol{u}\cdot\boldsymbol{v}}{\boldsymbol{v}\cdot\boldsymbol{v}}\right)\boldsymbol{v}$　　6. $\left(\dfrac{\boldsymbol{x}\cdot\boldsymbol{w}}{\boldsymbol{x}\cdot\boldsymbol{x}}\right)\boldsymbol{x}$　　7. $\|\boldsymbol{w}\|$　　8. $\|\boldsymbol{x}\|$

习题 9 ~ 12 中，要求给出一个与给定向量同向的单位向量.

9. $\begin{bmatrix}-30\\40\end{bmatrix}$　　10. $\begin{bmatrix}-6\\4\\-3\end{bmatrix}$　　11. $\begin{bmatrix}7/4\\1/2\\1\end{bmatrix}$　　12. $\begin{bmatrix}8/3\\2\end{bmatrix}$

13. 计算向量 $\boldsymbol{x}=\begin{bmatrix}10\\-3\end{bmatrix}$ 与 $\boldsymbol{y}=\begin{bmatrix}-1\\-5\end{bmatrix}$ 的距离.

14. 计算向量 $\boldsymbol{u}=\begin{bmatrix}0\\-5\\2\end{bmatrix}$ 与 $\boldsymbol{z}=\begin{bmatrix}-4\\-1\\8\end{bmatrix}$ 的距离.

判断习题 15 ~ 18 中的向量是否互相正交.

15. $\boldsymbol{a}=\begin{bmatrix}8\\-5\end{bmatrix}$　$\boldsymbol{b}=\begin{bmatrix}-2\\-3\end{bmatrix}$　　16. $\boldsymbol{u}=\begin{bmatrix}12\\3\\-5\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}2\\-3\\3\end{bmatrix}$

17. $\boldsymbol{u}=\begin{bmatrix}3\\2\\-5\\0\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}-4\\1\\-2\\6\end{bmatrix}$　　18. $\boldsymbol{y}=\begin{bmatrix}-3\\7\\4\\0\end{bmatrix}$，$\boldsymbol{z}=\begin{bmatrix}1\\-8\\15\\-7\end{bmatrix}$

习题 19 ~ 20 中，所有向量都属于 $\mathbf{R}^n$. 判断其中每个论述正确与否，并说明理由.

19. a. $\boldsymbol{v}\cdot\boldsymbol{v}=\|\boldsymbol{v}\|^2$.
    b. 对任意数量 c 都有 $\boldsymbol{u}\cdot(c\boldsymbol{v})=c(\boldsymbol{u}\cdot\boldsymbol{v})$.

c. 如果 $\boldsymbol{u}$ 到 $\boldsymbol{v}$ 的距离等于 $\boldsymbol{u}$ 到 $-\boldsymbol{v}$ 的距离，则 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 正交.

d. 对于方阵 $\boldsymbol{A}$，Col $\boldsymbol{A}$ 中的向量与 Nul $\boldsymbol{A}$ 中的向量正交.

e. 如果向量 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_p$ 张成一个子空间 W，并且 $\boldsymbol{x}$ 与 $\boldsymbol{v}_j, j=1,\cdots,p$ 都正交，则 $\boldsymbol{x}$ 属于 $W^{\perp}$.

20. a. $\boldsymbol{u}\cdot\boldsymbol{v}-\boldsymbol{v}\cdot\boldsymbol{u}=0$.

b. 对任意数量 c 都有 $\|c\boldsymbol{v}\|=c\|\boldsymbol{v}\|$.

382

c. 如果 $\boldsymbol{x}$ 与某个子空间 W 中的每一向量都正交，则 $\boldsymbol{x}$ 属于 $W^{\perp}$.

d. 如果 $\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2=\|\boldsymbol{u}+\boldsymbol{v}\|^2$，则 $\boldsymbol{u}$ 与 $\boldsymbol{v}$ 正交.

e. 对任意 $m\times n$ 矩阵 $\boldsymbol{A}$，$\boldsymbol{A}$ 的解空间中的向量与 $\boldsymbol{A}$ 的行向量空间中的向量正交.

21. 利用内积的转置定义，证明定理1的结论(b)和(c). 可以引用第2章相关事实.

22. 令 $\boldsymbol{u}=(u_1,u_2,u_3)$，说明为什么 $\boldsymbol{u}\cdot\boldsymbol{u}\geqslant 0$ 以及何时有 $\boldsymbol{u}\cdot\boldsymbol{u}=0$.

23. 令 $\boldsymbol{u}=\begin{bmatrix}2\\-5\\-1\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}-7\\-4\\6\end{bmatrix}$. 不用勾股定理，计算并且比较 $\boldsymbol{u}\cdot\boldsymbol{v}$，$\|\boldsymbol{u}\|^2$，$\|\boldsymbol{v}\|^2$ 和 $\|\boldsymbol{u}+\boldsymbol{v}\|^2$ 的值.

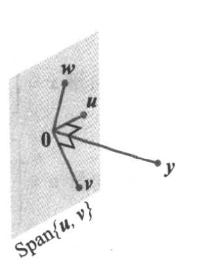
24. 对 $\mathbf{R}^n$ 中向量 $\boldsymbol{u}$，$\boldsymbol{v}$ 验证平行四边形法则：$\|\boldsymbol{u}+\boldsymbol{v}\|^2+\|\boldsymbol{u}-\boldsymbol{v}\|^2=2\|\boldsymbol{u}\|^2+2\|\boldsymbol{v}\|^2$.

25. 令 $\boldsymbol{v}=\begin{bmatrix}a\\b\end{bmatrix}$，写出由全体与 $\boldsymbol{v}$ 正交的向量 $\begin{bmatrix}x\\y\end{bmatrix}$ 构成的集合 H. [提示：分 $\boldsymbol{v}=\mathbf{0}$ 和 $\boldsymbol{v}\neq\mathbf{0}$ 两种情况考虑.]

26. 令 $\boldsymbol{u}=\begin{bmatrix}5\\-6\\7\end{bmatrix}$，$W$ 是 $\mathbf{R}^3$ 中全体满足 $\boldsymbol{u}\cdot\boldsymbol{x}=0$ 的 $\boldsymbol{x}$ 所构成的集合. 第4章哪个定理表明 W 是 $\mathbf{R}^3$ 的子空间？试用几何语言描述 W.

27. 假设向量 $\boldsymbol{y}$ 与向量 $\boldsymbol{u}$，$\boldsymbol{v}$ 正交，证明 $\boldsymbol{y}$ 与向量 $\boldsymbol{u}+\boldsymbol{v}$ 正交.

28. 假设向量 $\boldsymbol{y}$ 与 $\boldsymbol{u}$，$\boldsymbol{v}$ 正交，证明 $\boldsymbol{y}$ 与 Span$\{\boldsymbol{u},\boldsymbol{v}\}$中的每一向量 $\boldsymbol{w}$ 正交. [提示：Span$\{\boldsymbol{u},\boldsymbol{v}\}$中的任意向量 $\boldsymbol{w}$ 都可以写成 $\boldsymbol{w}=c_1\boldsymbol{u}+c_2\boldsymbol{v}$，因此只需证明 $\boldsymbol{y}$ 与这种形式的向量 $\boldsymbol{w}$ 正交]



29. 令 $W=\text{Span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_p\}$，证明如果 $\boldsymbol{x}$ 与 $\boldsymbol{v}_j$，$j=1,\cdots,p$ 都正交，则 $\boldsymbol{x}$ 与 W 中的每个向量都正交.

30. 令 W 是 $\mathbf{R}^n$ 的子空间，$W^{\perp}$ 是与 W 正交的全体向量所构成的集合. 按照下列步骤证明 $W^{\perp}$ 是 $\mathbf{R}^n$ 的子空间：

a. 任取 $\boldsymbol{z}$ 属于 $W^{\perp}$，$\boldsymbol{u}$ 属于 W，则 $\boldsymbol{z}\cdot\boldsymbol{u}=0$. 任取数量 c，证明 $c\boldsymbol{z}$ 与 $\boldsymbol{u}$ 正交(由 $\boldsymbol{u}$ 的任意性可知 $c\boldsymbol{z}$ 属于 $W^{\perp}$).

b. 任取 $\boldsymbol{z}_1$，$\boldsymbol{z}_2$ 属于 $W^{\perp}$，$\boldsymbol{u}$ 属于 W，证明 $\boldsymbol{z}_1+\boldsymbol{z}_2$ 与 $\boldsymbol{u}$ 正交. 有关 $\boldsymbol{z}_1+\boldsymbol{z}_2$ 可以做出什么结论？为什么？

c. 结束证明，得出结论：$W^{\perp}$ 是 $\mathbf{R}^n$ 的子空间.

31. 证明：若 $\boldsymbol{x}$ 同时属于 W 和 $W^{\perp}$，则 $\boldsymbol{x}=\mathbf{0}$.

32. [M] 构造 $\mathbf{R}^4$ 的一对随机向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$，令

$$\boldsymbol{A}=\begin{bmatrix}0.5 & 0.5 & 0.5 & 0.5\\0.5 & 0.5 & -0.5 & -0.5\\0.5 & -0.5 & 0.5 & -0.5\\0.5 & -0.5 & -0.5 & 0.5\end{bmatrix}$$

a. 将 A 的列向量分别记作 $a_1,\cdots,a_4$. 计算每一列向量的长度以及 $a_1 \cdot a_2$, $a_1 \cdot a_3$, $a_1 \cdot a_4$, $a_2 \cdot a_3$, $a_2 \cdot a_4$ 和 $a_3 \cdot a_4$ 的值.

b. 计算并比较 u, Au, v 和 Av 的长度.

c. 利用本节公式(2)计算 u, v 夹角的余弦值，将它与 Au, Av 夹角的余弦值做比较.

d. 另选两对向量，重复(b)，(c)的计算. 关于矩阵 A 对向量作用的影响，你能做何猜想?

33. [M] 生成 $\mathbf{R}^4$ 的整系数随机向量 x, y 及 $v(v\neq 0)$，计算以下各量：

$$\left(\frac{x\cdot v}{v\cdot v}\right)v,\left(\frac{y\cdot v}{v\cdot v}\right)v,\frac{(x+y)\cdot v}{v\cdot v}v,\frac{(10x)\cdot v}{v\cdot v}v$$

另选随机向量 x 和 y，重复上述计算. 关于映射 $x\to T(x)=\left(\frac{x\cdot v}{v\cdot v}\right)v\,(v\neq 0)$，你能做何猜想? 使用代数方法验证你的猜想.

34. [M] 令 $A=\begin{bmatrix}-6 & 3 & -27 & -33 & -13\\ 6 & -5 & 25 & 28 & 14\\ 8 & -6 & 34 & 38 & 18\\ 12 & -10 & 50 & 41 & 23\\ 14 & -21 & 49 & 29 & 33\end{bmatrix}$.

构造矩阵 N 和 $\mathbf{R}$，使得 N 的列向量构成 Nul A 的一组基，$\mathbf{R}$ 的行向量构成 Row A 的一组基(细节可参考 4.6 节). 利用 N 和 $\mathbf{R}$ 进行计算，验证定理 3 所描述的事实. 383

基础练习答案

1. $a\cdot b=7$, $a\cdot a=5$. 于是 $\frac{a\cdot b}{a\cdot a}=\frac{7}{5}$, $\left(\frac{a\cdot b}{a\cdot a}\right)a=\frac{7}{5}a=\begin{bmatrix}-14/5\\ 7/5\end{bmatrix}$.

2. 适当缩放 c，乘以 3 得到 $y=\begin{bmatrix}4\\ -3\\ 2\end{bmatrix}$. 计算得 $\|y\|^2=29$, $\|y\|=\sqrt{29}$. 于是与 c, y 同向的单位向量是 $u=\frac{1}{\|y\|}y=\begin{bmatrix}4/\sqrt{29}\\ -3/\sqrt{29}\\ 2/\sqrt{29}\end{bmatrix}$. 3. 由于

$$d\cdot c=\begin{bmatrix}5\\ 6\\ -1\end{bmatrix}\cdot\begin{bmatrix}4/3\\ -1\\ 2/3\end{bmatrix}=\frac{20}{3}-6-\frac{2}{3}=0$$

故 d 正交于 c.

4. 由于存在某个数量 k,使得 u 等于 kc,且

$$d\cdot u=d\cdot(kc)=k(d\cdot c)=k(0)=0$$

故 d 正交于 u.

6.2 正 交 集

设 $\{u_1,\cdots,u_p\}$ 是 $\mathbf{R}^n$ 中一个向量集合，如果其中任意两个相异向量都互相正交，即 $i\neq j$ 时有 $u_i\cdot u_j=0$，则称这个集合是**正交集**(orthogonal set).

【例题 1】 证明 $\{u_1,u_2,u_3\}$ 是正交集，其中，

$$u_1=\begin{bmatrix}3\\1\\1\end{bmatrix},\ u_2=\begin{bmatrix}-1\\2\\1\end{bmatrix},\ u_3=\begin{bmatrix}-1/2\\-2\\7/2\end{bmatrix}$$

解： 考虑不同向量组成向量对的三种可能，即，$\{u_1,u_2\}$，$\{u_1,u_3\}$和$\{u_2,u_3\}$.

$$u_1\cdot u_2=3(-1)+1(2)+1(1)=0$$

$$u_1\cdot u_3=3\left(-\frac{1}{2}\right)+1(-2)+1\left(\frac{7}{2}\right)=0$$

$$u_2\cdot u_3=-1\left(-\frac{1}{2}\right)+2(-2)+1\left(\frac{7}{2}\right)=0$$

每一对相异的向量都正交，因此$\{u_1,u_2,u_3\}$是正交集．如图6-11，这三条线段两两垂直．■

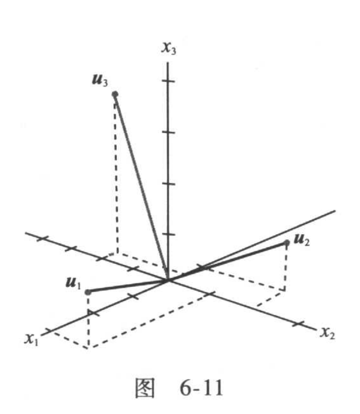


图　6-11

384

【定理4】　若$S=\{u_1,\cdots,u_p\}$是$\mathbf{R}^n$中非零向量构成的正交集，则S线性无关且S是它所张成的子空间的一组基.

证明： 若存在数量$c_1,\cdots,c_p$使得$\mathbf{0}=c_1u_1+\cdots+c_pu_p$，由$u_1$正交于$u_2,\cdots,u_p$可得：

$$\begin{aligned}0&=\mathbf{0}\cdot u_1=(c_1u_1+c_2u_2+\cdots+c_pu_p)\cdot u_1\\&=(c_1u_1)\cdot u_1+(c_2u_2)\cdot u_1+\cdots+(c_pu_p)\cdot u_1\\&=c_1(u_1\cdot u_1)+c_2(u_2\cdot u_1)+\cdots+c_p(u_p\cdot u_1)\\&=c_1(u_1\cdot u_1)\end{aligned}$$

又因为u_1是非零向量，$u_1\cdot u_1$不是零，所以$c_1=0$．类似地，c_2，…，c_p必须全为零．因此S是线性无关的．■

【定义】　$\mathbf{R}^n$中子空间W的一组**正交基**(orthogonal basis)是W的一组基并且本身构成正交集.

下面的定理说明正交基比其他基要好：线性组合的权重更容易求得.

【定理5】　令$\{u_1,\cdots,u_p\}$是$\mathbf{R}^n$的子空间W的一组正交基．对任意$y\in W$，线性组合

$$y=c_1u_1+\cdots+c_pu_p$$

的权重可由下式给出：

$$c_j=\frac{y\cdot u_j}{u_j\cdot u_j}\quad(j=1,\cdots,p)$$

证明： 与上一定理的证明相同，$\{u_1,\cdots,u_p\}$的正交性表明：

$$y\cdot u_1=(c_1u_1+c_2u_2+\cdots+c_pu_p)\cdot u_1=c_1(u_1\cdot u_1)$$

由于$u_1\cdot u_1$不是零，上述等式可以解出c_1，为了求得$c_j(j=1,\cdots,p)$，只需计算$y\cdot u_j$再求解c_j.■

【例题2】　例题1中的集合$S=\{u_1,u_2,u_3\}$是$\mathbf{R}^n$的一组正交基．将向量$y=\begin{bmatrix}6\\1\\-8\end{bmatrix}$表示成$S$中向量的线性组合.

解：计算：

$$\boldsymbol{y}\cdot\boldsymbol{u}_1=11,\qquad \boldsymbol{y}\cdot\boldsymbol{u}_2=-12,\qquad \boldsymbol{y}\cdot\boldsymbol{u}_3=-33$$

$$\boldsymbol{u}_1\cdot\boldsymbol{u}_1=11,\qquad \boldsymbol{u}_2\cdot\boldsymbol{u}_2=6,\qquad \boldsymbol{u}_3\cdot\boldsymbol{u}_3=33/2$$

385

根据定理5，有

$$\boldsymbol{y}=\frac{\boldsymbol{y}\cdot\boldsymbol{u}_1}{\boldsymbol{u}_1\cdot\boldsymbol{u}_1}\boldsymbol{u}_1+\frac{\boldsymbol{y}\cdot\boldsymbol{u}_2}{\boldsymbol{u}_2\cdot\boldsymbol{u}_2}\boldsymbol{u}_2+\frac{\boldsymbol{y}\cdot\boldsymbol{u}_3}{\boldsymbol{u}_3\cdot\boldsymbol{u}_3}\boldsymbol{u}_3$$

$$=\frac{11}{11}\boldsymbol{u}_1+\frac{-12}{6}\boldsymbol{u}_2+\frac{-33}{33/2}\boldsymbol{u}_3=\boldsymbol{u}_1-2\boldsymbol{u}_2-2\boldsymbol{u}_3$$ ■

注意，要计算 $\boldsymbol{y}$ 关于一组正交基的线性表示，是非常容易的. 但是，如第1章所述，如果不是正交基，为了求出权重就需要求解线性方程组.

下面我们来讨论一种构造，它在涉及正交性的很多计算中非常有用，并且将会给出定理5的一种几何意义上的解释.

6.2.1 正交投影

给定 $\mathbf{R}^n$ 中一个非零向量 $\boldsymbol{u}$，考虑如何将 $\mathbf{R}^n$ 中的向量 $\boldsymbol{y}$ 分解成两个向量之和，其中一个分量是 $\boldsymbol{u}$ 的倍数，另一个分量与 $\boldsymbol{u}$ 正交. 我们希望写成：

$$\boldsymbol{y}=\hat{\boldsymbol{y}}+\boldsymbol{z} \tag{1}$$

使得对某一数量 α 有 $\hat{\boldsymbol{y}}=\alpha\boldsymbol{u}$，并且 $\boldsymbol{z}$ 是与 $\boldsymbol{u}$ 正交的某个向量. 如图6-12所示. 任意给定一个数量 α，令 $\boldsymbol{z}=\boldsymbol{y}-\alpha\boldsymbol{u}$，则(1)式被满足. $\boldsymbol{y}-\hat{\boldsymbol{y}}$ 与 $\boldsymbol{u}$ 正交当且仅当

$$0=(\boldsymbol{y}-\alpha\boldsymbol{u})\cdot\boldsymbol{u}=\boldsymbol{y}\cdot\boldsymbol{u}-(\alpha\boldsymbol{u})\cdot\boldsymbol{u}=\boldsymbol{y}\cdot\boldsymbol{u}-\alpha(\boldsymbol{u}\cdot\boldsymbol{u})$$

即，(1)式被正交于 $\boldsymbol{u}$ 的向量 $\boldsymbol{z}$ 满足当且仅当 $\alpha=\dfrac{\boldsymbol{y}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}}$ 及 $\hat{\boldsymbol{y}}=\dfrac{\boldsymbol{y}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}}\boldsymbol{u}$. 这个向量 $\hat{\boldsymbol{y}}$ 称作 **$\boldsymbol{y}$ 在 $\boldsymbol{u}$ 上的正交投影**(orthogonal projection of $\boldsymbol{y}$ onto $\boldsymbol{u}$)，向量 $\boldsymbol{z}$ 则称作 **$\boldsymbol{y}$ 中正交于 $\boldsymbol{u}$ 的分量**(component of $\boldsymbol{y}$ orthogonal to $\boldsymbol{u}$).

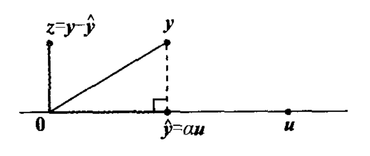


图6-12 求 α，使得 $\boldsymbol{y}-\hat{\boldsymbol{y}}$ 与 $\boldsymbol{u}$ 正交

如果 c 是一个非零数量，在 $\hat{\boldsymbol{y}}$ 的定义中以 $c\boldsymbol{u}$ 代替 $\boldsymbol{u}$，则 $\boldsymbol{y}$ 在 $c\boldsymbol{u}$ 上的正交投影刚好等于 $\boldsymbol{y}$ 在 $\boldsymbol{u}$ 上的正交投影(习题31). 因此这种投影由 $\boldsymbol{u}$ 所张成的子空间 L(经过 $\boldsymbol{u}$ 和 $\boldsymbol{0}$ 的直线)唯一确定. $\hat{\boldsymbol{y}}$ 有时也记作 $\mathrm{proj}_L\,\boldsymbol{y}$，并称作 **$\boldsymbol{y}$ 在 L 上的正交投影**(orthogonal projection of $\boldsymbol{y}$ onto L). 即

$$\boxed{\hat{\boldsymbol{y}}=\mathrm{proj}_L\,\boldsymbol{y}=\frac{\boldsymbol{y}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}}\boldsymbol{u}} \tag{2}$$

【例题3】 令 $\boldsymbol{y}=\begin{bmatrix}7\\6\end{bmatrix}$，$\boldsymbol{u}=\begin{bmatrix}4\\2\end{bmatrix}$. 求 $\boldsymbol{y}$ 在 $\boldsymbol{u}$ 上的正交投影. 然后将 $\boldsymbol{y}$ 写成互相正交的两个向量之和，其中一个属于 Span$\{\boldsymbol{u}\}$，另一个与 $\boldsymbol{u}$ 正交.

386

解：计算

$$\boldsymbol{y}\cdot\boldsymbol{u}=\begin{bmatrix}7\\6\end{bmatrix}\cdot\begin{bmatrix}4\\2\end{bmatrix}=40$$

$$\boldsymbol{u}\cdot\boldsymbol{u}=\begin{bmatrix}4\\2\end{bmatrix}\cdot\begin{bmatrix}4\\2\end{bmatrix}=20$$

y 在 u 上的正交投影是：

$$\hat{y}=\frac{y\cdot u}{u\cdot u}u=\frac{40}{20}u=2\begin{bmatrix}4\\2\end{bmatrix}=\begin{bmatrix}8\\4\end{bmatrix}$$

y 中正交于 u 的分量是：

$$y-\hat{y}=\begin{bmatrix}7\\6\end{bmatrix}-\begin{bmatrix}8\\4\end{bmatrix}=\begin{bmatrix}-1\\2\end{bmatrix}$$

这两个向量之和是 y，即，

$$\underset{\substack{\uparrow\\ y}}{\begin{bmatrix}7\\6\end{bmatrix}}=\underset{\substack{\uparrow\\ \hat{y}}}{\begin{bmatrix}8\\4\end{bmatrix}}+\underset{\substack{\uparrow\\ (y-\hat{y})}}{\begin{bmatrix}-1\\2\end{bmatrix}}$$

图 6-13 描述了 y 的这一分解．注意：如果上述计算是正确的，则 $\{\hat{y},\ y-\hat{y}\}$ 构成一个正交集．为验证这一点，只需计算：

$$\hat{y}\cdot(y-\hat{y})=\begin{bmatrix}8\\4\end{bmatrix}\cdot\begin{bmatrix}-1\\2\end{bmatrix}=-8+8=0$$ ■

根据 $\hat{y}$ 的构造，图 6-13 中 y 与 $\hat{y}$ 之间的线段垂直于 L，因此 $\hat{y}$ 所确定的那一点是 L 上与 y 距离最近的点．（可以从几何角度出发证明这一点，我们现在假定它对于 $\mathbf{R}^2$ 是成立的，在 6.3 节中将证明 $\mathbf{R}^n$ 的情形．）

387 **【例题 4】**　求图 6-13 中 y 到 L 的距离．

解： y 到 L 的距离是从 y 到正交投影 $\hat{y}$ 之间垂直线段的长度．这个长度等于 $y-\hat{y}$ 的长度．于是所求的距离是：

$$\|y-\hat{y}\|=\sqrt{(-1)^2+2^2}=\sqrt{5}$$ ■

图 6-13　y 在通过原点的直线 L 上的正交投影

6.2.2　定理 5 的几何解释

(2) 式所给出的正交投影 $\hat{y}$ 的表达式与定理 5 中线性组合的每一项形式相同．因此，定理 5 将向量 y 分解成若干向量之和，其中每一分量都是 y 在一个 1 维子空间上的正交投影．

可以很容易地想象，$W=\mathbf{R}^2=\mathrm{Span}\{u_1,u_2\}$，且 u_1 和 u_2 正交的情形．任意 $y\in\mathbf{R}^2$，y 可以写成如下形式：

$$y = \frac{y \cdot u_1}{u_1 \cdot u_1}u_1 + \frac{y \cdot u_2}{u_2 \cdot u_2}u_2 \tag{3}$$

(3)式中的第一项是 y 在 u_1 所张成子空间(即通过 u_1 和原点的直线)上的投影，第二项是 y 在 u_2 所张成子空间上的投影. 于是(3)式将 y 表示成它在由 u_1 和 u_2 所确定的(正交)坐标轴上的投影之和. 如图 6-14 所示.

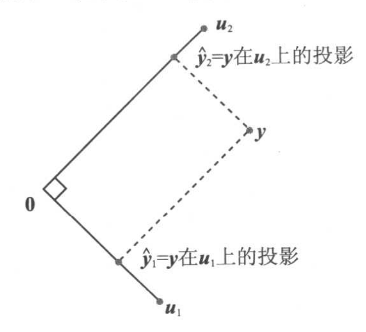


图 6-14 将一个向量分解成两个投影向量之和

定理 5 将 Span$\{u_1, \cdots, u_p\}$中的每一向量 y 分解成 p 个 1 维子空间上投影的和，并且这些投影互相正交.

6.2.3 力的分解

物理学中，向一个物体施加某种力时，可能需要用到图 6-14 所示的分解. 选择一个适当的坐标系以后，我们便可以用 $\mathbf{R}^2$ 或 $\mathbf{R}^3$ 中的向量 y 来表示力的作用. 通常待解决的问题中包含一个重要而特殊的方向，记作向量 u. 例如，如果对一个物体施加某个作用力使其做直线运动，则向量 u 应该表示运动的方向，如图 6-15 所示. 解决问题的关键是将作用力分解成为与 u 同向以及与 u 正交的两个分量之和. 计算过程与前述例题 3 类似.

388

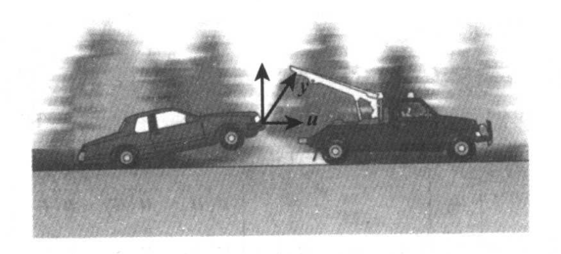


图 6-15

6.2.4 标准正交集

如果集合$\{u_1, \cdots, u_p\}$是单位向量所构成的正交集，则称之为**标准正交集**(orthonormal set). 如果 W 是由标准正交集$\{u_1, \cdots, u_p\}$张成的子空间，由定理 4 可知$\{u_1, \cdots, u_p\}$线性无关，因此将它称作 W 的一组**标准正交基**(orthonormal basis).

标准正交集最简单的例子是 $\mathbf{R}^n$ 的标准基$\{e_1, \cdots, e_n\}$. $\{e_1, \cdots, e_n\}$的任意非空子集

也是标准正交的. 下面给出了一个较为复杂的例子.

【例题5】　证明$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$是$\mathbf{R}^3$的一组标准正交基，其中

$$\boldsymbol{v}_1=\begin{bmatrix}3/\sqrt{11}\\1/\sqrt{11}\\1/\sqrt{11}\end{bmatrix},\quad \boldsymbol{v}_2=\begin{bmatrix}-1/\sqrt{6}\\2/\sqrt{6}\\1/\sqrt{6}\end{bmatrix},\quad \boldsymbol{v}_3=\begin{bmatrix}-1/\sqrt{66}\\-4/\sqrt{66}\\7/\sqrt{66}\end{bmatrix}$$

解： 计算

$$\boldsymbol{v}_1\cdot\boldsymbol{v}_2=-3/\sqrt{66}+2/\sqrt{66}+1/\sqrt{66}=0$$
$$\boldsymbol{v}_1\cdot\boldsymbol{v}_3=-3/\sqrt{726}-4/\sqrt{726}+7/\sqrt{726}=0$$
$$\boldsymbol{v}_2\cdot\boldsymbol{v}_3=1/\sqrt{396}-8/\sqrt{396}+7/\sqrt{396}=0$$

这样$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$是正交集，此外还有，

$$\boldsymbol{v}_1\cdot\boldsymbol{v}_1=9/11+1/11+1/11=1$$
$$\boldsymbol{v}_2\cdot\boldsymbol{v}_2=1/6+4/6+1/6=1$$
$$\boldsymbol{v}_3\cdot\boldsymbol{v}_3=1/66+16/66+49/66=1$$

389

这说明$\boldsymbol{v}_1$，$\boldsymbol{v}_2$和$\boldsymbol{v}_3$都是单位向量. 因此$\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$是一个标准正交集. 又因为这个集合线性无关，其中的三个向量构成$\mathbf{R}^3$的一组基. 如图6-16所示. ■

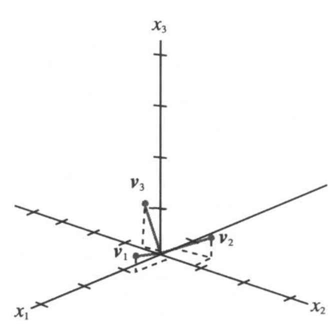


图6-16　$\mathbf{R}^3$的一个标准正交基

对于一个由非零向量组成的正交集，如果将其中的向量标准化使之具有单位长度，那么得到的新向量仍然正交，于是得到的新集合将会是一个标准正交集. 见习题32. 容易验证图6-16(例题5)中的向量是与图6-11(例题1)中向量同向的单位向量.

在实际应用以及矩阵运算的计算机算法当中，列向量组成标准正交集的那些矩阵都十分重要. 定理6和7给出了这类矩阵的主要性质.

【定理6】　一个$m\times n$矩阵U有标准正交的列向量当且仅当$U^{\mathrm{T}}U=I$.

证明： 为简化记号，我们假定U只包含三列，其中每一列都是$\mathbf{R}^m$中的一个向量. 一般情形的证明本质上都与此相同. 令$U=[\boldsymbol{u}_1\quad\boldsymbol{u}_2\quad\boldsymbol{u}_3]$并计算：

$$U^{\mathrm{T}}U=\begin{bmatrix}\boldsymbol{u}_1^{\mathrm{T}}\\\boldsymbol{u}_2^{\mathrm{T}}\\\boldsymbol{u}_3^{\mathrm{T}}\end{bmatrix}[\boldsymbol{u}_1\quad\boldsymbol{u}_2\quad\boldsymbol{u}_3]=\begin{bmatrix}\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{u}_1&\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{u}_2&\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{u}_3\\\boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{u}_1&\boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{u}_2&\boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{u}_3\\\boldsymbol{u}_3^{\mathrm{T}}\boldsymbol{u}_1&\boldsymbol{u}_3^{\mathrm{T}}\boldsymbol{u}_2&\boldsymbol{u}_3^{\mathrm{T}}\boldsymbol{u}_3\end{bmatrix}\tag{4}$$

右边矩阵中的元素是用转置记号所表示的内积. U的各列互相正交当且仅当

$$\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{u}_2=\boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{u}_1=0,\ \boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{u}_3=\boldsymbol{u}_3^{\mathrm{T}}\boldsymbol{u}_1=0,\ \boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{u}_3=\boldsymbol{u}_3^{\mathrm{T}}\boldsymbol{u}_2=0\tag{5}$$

U的各列有单位长度当且仅当

$$\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{u}_1=1,\ \boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{u}_2=1,\ \boldsymbol{u}_3^{\mathrm{T}}\boldsymbol{u}_3=1\tag{6}$$

由(4)~(6)即知，定理得证. ■

【定理7】 令 U 是一个 $m\times n$ 矩阵且有标准正交的列向量，令 x，$y\in\mathbf{R}^n$，则：

(a) $\|Ux\|=\|x\|$

(b) $(Ux)\cdot(Uy)=x\cdot y$

(c) $(Ux)\cdot(Uy)=0$ 当且仅当 $x\cdot y=0$

性质(a)和(c)表明线性映射 $x\mapsto Ux$ 保持长度和正交性．这些性质对许多计算机算法都至关重要．定理7的证明可参考习题25.

390

【例题6】 令 $U=\begin{bmatrix}1/\sqrt{2} & 2/3\\ 1/\sqrt{2} & -2/3\\ 0 & 1/3\end{bmatrix}$，$x=\begin{bmatrix}\sqrt{2}\\ 3\end{bmatrix}$．注意到 U 有标准正交的列向量，且

$$U^{\mathrm{T}}U=\begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2} & 0\\ 2/3 & -2/3 & 1/3\end{bmatrix}\begin{bmatrix}1/\sqrt{2} & 2/3\\ 1/\sqrt{2} & -2/3\\ 0 & 1/3\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

验证 $\|Ux\|=\|x\|$.

解：

$$Ux=\begin{bmatrix}1/\sqrt{2} & 2/3\\ 1/\sqrt{2} & -2/3\\ 0 & 1/3\end{bmatrix}\begin{bmatrix}\sqrt{2}\\ 3\end{bmatrix}=\begin{bmatrix}3\\ -1\\ 1\end{bmatrix}$$

$$\|Ux\|=\sqrt{9+1+1}=\sqrt{11}$$

$$\|x\|=\sqrt{2+9}=\sqrt{11}$$ ■

定理6和7在处理方阵时尤其有用．一个**正交矩阵**(orthogonal matrix)是满足条件 $U^{-1}=U^{\mathrm{T}}$ 的可逆方阵 U．根据定理6，这类矩阵有标准正交的列向量.[1] 显然，有标准正交列向量的方阵是正交矩阵．出人意料的是，这类矩阵也有标准正交的行向量．见习题27和28．第7章中我们将经常使用正交矩阵.

【例题7】 矩阵 $U=\begin{bmatrix}3/\sqrt{11} & -1/\sqrt{6} & -1/\sqrt{66}\\ 1/\sqrt{11} & 2/\sqrt{6} & -4/\sqrt{66}\\ 1/\sqrt{11} & 1/\sqrt{6} & 7/\sqrt{66}\end{bmatrix}$ 是正交矩阵，这是因为 U 是方阵并且根据例题5，U 中各列标准正交．可以验证 U 中各行也是标准正交的. ■

基础练习

1. 令 $u_1=\begin{bmatrix}-1/\sqrt{5}\\ 2/\sqrt{5}\end{bmatrix}$，$u_2=\begin{bmatrix}2/\sqrt{5}\\ 1/\sqrt{5}\end{bmatrix}$．证明 $\{u_1,u_2\}$ 是 $\mathbf{R}^2$ 的正交基.

---

1. 称作标准正交矩阵可能更好，一些统计学教材里采用了这个术语．但在线性代数中，标准的术语是正交矩阵.

2. 令 y 和 L 如例题3和图6-13. 以 $u=\begin{bmatrix}2\\1\end{bmatrix}$ 代替例题3中的 u，计算 y 在 L 上的正交投影 $\hat{y}$.

391 3. 令 U 和 x 如例题6，令 $y=\begin{bmatrix}-3\sqrt{2}\\6\end{bmatrix}$. 验证 $Ux\cdot Uy=x\cdot y$.

习题　6.2

判断习题1～6中，哪些向量集合是正交的.

1. $\begin{bmatrix}-1\\4\\-3\end{bmatrix},\begin{bmatrix}5\\2\\1\end{bmatrix},\begin{bmatrix}3\\-4\\-7\end{bmatrix}$　2. $\begin{bmatrix}1\\-2\\1\end{bmatrix},\begin{bmatrix}0\\1\\2\end{bmatrix},\begin{bmatrix}-5\\-2\\1\end{bmatrix}$　3. $\begin{bmatrix}2\\-7\\-1\end{bmatrix},\begin{bmatrix}-6\\-3\\9\end{bmatrix},\begin{bmatrix}3\\1\\-1\end{bmatrix}$

4. $\begin{bmatrix}2\\-5\\-3\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}4\\-2\\6\end{bmatrix}$　5. $\begin{bmatrix}3\\-2\\1\\3\end{bmatrix},\begin{bmatrix}-1\\3\\-3\\4\end{bmatrix},\begin{bmatrix}3\\8\\7\\0\end{bmatrix}$　6. $\begin{bmatrix}5\\-4\\0\\3\end{bmatrix},\begin{bmatrix}-4\\1\\-3\\8\end{bmatrix},\begin{bmatrix}3\\3\\5\\-1\end{bmatrix}$

证明习题7～10中，$\{u_1,u_2\}$ 或 $\{u_1,u_2,u_3\}$ 是 $\mathbf{R}^2$ 或 $\mathbf{R}^3$ 的正交基，然后将 x 表示成这些向量 u 的线性组合.

7. $u_1=\begin{bmatrix}2\\-3\end{bmatrix}$, $u_2=\begin{bmatrix}6\\4\end{bmatrix}$, $x=\begin{bmatrix}9\\-7\end{bmatrix}$

8. $u_1=\begin{bmatrix}3\\1\end{bmatrix}$, $u_2=\begin{bmatrix}-2\\6\end{bmatrix}$, $x=\begin{bmatrix}-6\\3\end{bmatrix}$

9. $u_1=\begin{bmatrix}1\\0\\1\end{bmatrix}$, $u_2=\begin{bmatrix}-1\\4\\1\end{bmatrix}$, $u_3=\begin{bmatrix}2\\1\\-2\end{bmatrix}$ 且 $x=\begin{bmatrix}8\\-4\\-3\end{bmatrix}$

10. $u_1=\begin{bmatrix}3\\-3\\0\end{bmatrix}$, $u_2=\begin{bmatrix}2\\2\\-1\end{bmatrix}$, $u_3=\begin{bmatrix}1\\1\\4\end{bmatrix}$ 且 $x=\begin{bmatrix}5\\-3\\1\end{bmatrix}$

11. 计算 $\begin{bmatrix}1\\7\end{bmatrix}$ 在通过 $\begin{bmatrix}-4\\2\end{bmatrix}$ 与原点的直线上的正交投影.

12. 计算 $\begin{bmatrix}1\\-1\end{bmatrix}$ 在通过 $\begin{bmatrix}-1\\3\end{bmatrix}$ 与原点的直线上的正交投影.

13. 令 $y=\begin{bmatrix}2\\3\end{bmatrix}$, $u=\begin{bmatrix}4\\-7\end{bmatrix}$. 将 y 写成互相正交的两向量之和，一个属于 Span$\{u\}$，另一个与 u 正交.

14. 令 $y=\begin{bmatrix}2\\6\end{bmatrix}$, $u=\begin{bmatrix}7\\1\end{bmatrix}$. 将 y 写成互相正交的两向量之和，一个属于 Span$\{u\}$，另一个与 u

正交.

15. 令 $y=\begin{bmatrix}3\\1\end{bmatrix}$, $u=\begin{bmatrix}8\\6\end{bmatrix}$. 计算 y 到过 u 及原点的直线的距离.

16. 令 $y=\begin{bmatrix}-3\\9\end{bmatrix}$, $u=\begin{bmatrix}1\\2\end{bmatrix}$. 计算 y 到过 u 及原点的直线的距离.

判断习题 17 ~ 22 中，哪些向量集合是标准正交的. 如果向量集仅仅是正交的，将其中向量标准化以得到标准正交集.

17. $\begin{bmatrix}1/3\\1/3\\1/3\end{bmatrix}, \begin{bmatrix}-1/2\\0\\1/2\end{bmatrix}$ 18. $\begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\-1\\0\end{bmatrix}$ 19. $\begin{bmatrix}-0.6\\0.8\end{bmatrix}, \begin{bmatrix}0.8\\0.6\end{bmatrix}$ 20. $\begin{bmatrix}-2/3\\1/3\\2/3\end{bmatrix}, \begin{bmatrix}1/3\\2/3\\0\end{bmatrix}$

21. $\begin{bmatrix}1/\sqrt{10}\\3/\sqrt{20}\\3/\sqrt{20}\end{bmatrix}, \begin{bmatrix}3/\sqrt{10}\\-1/\sqrt{20}\\-1/\sqrt{20}\end{bmatrix}, \begin{bmatrix}0\\-1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$ 22. $\begin{bmatrix}1/\sqrt{18}\\4/\sqrt{18}\\1/\sqrt{18}\end{bmatrix}, \begin{bmatrix}1/\sqrt{2}\\0\\-1/\sqrt{2}\end{bmatrix}, \begin{bmatrix}-2/3\\1/3\\-2/3\end{bmatrix}$

习题 23 和 24 中的全体向量都属于 $\mathbf{R}^n$. 判断各命题的真假，并说明理由.

23. a. 并非 $\mathbf{R}^n$ 中的每个线性无关集都是正交集.
    b. 若 y 是某个正交集中非零向量的线性组合，则无需对矩阵施行行变换就能计算出线性组合的权重.
    c. 对由非零向量构成的正交集中的向量进行标准化，所得的新向量可能不正交.
    d. 具有标准正交列的矩阵是正交矩阵.
    e. 若 L 是通过 $\mathbf{0}$ 的直线且 $\hat{y}$ 是 y 在 L 上的正交投影，则 $\|\hat{y}\|$ 给出了 y 到 L 的距离.
24. a. 并非 $\mathbf{R}^n$ 中的每个正交集都是线性无关的.
    b. 若集合 $S=\{u_1,\cdots,u_p\}$ 具有性质：只要 $i\neq j$ 就有 $u_i\cdot u_j=0$，则 S 是一个标准正交集.
    c. 若 $m\times n$ 矩阵 A 的各列标准正交，则线性映射 $x\mapsto Ax$ 保持长度.
    d. 当 $c\neq 0$ 时，y 在 v 上的正交投影与 y 在 cv 上的正交投影相等.
    e. 正交矩阵可逆.

392

25. 证明定理 7. [提示：为证明(a)，可以计算 $\|Ux\|^2$ 或者先证明(b).]
26. 设 W 是 $\mathbf{R}^n$ 的子空间，且由 n 个非零正交向量张成. 说明为什么有 $W=\mathbf{R}^n$.
27. 令 U 是有标准正交列的方阵，说明 U 为什么可逆.（指出你所使用的定理）
28. 令 U 是 $n\times n$ 正交矩阵. 证明 U 的行向量构成 $\mathbf{R}^n$ 的一组标准正交基.
29. 令 U 和 V 是 $n\times n$ 正交矩阵. 说明 UV 为什么是正交矩阵. [即说明 UV 为什么可逆并且逆为 $(UV)^T$.]
30. 令 U 是 $n\times n$ 正交矩阵，V 是交换 U 中若干列而得到的矩阵. 说明 V 为什么是正交矩阵.
31. 证明：向量 y 在 $\mathbf{R}^2$ 中过原点的直线 L 上的正交投影不依赖于 $\hat{y}$ 的计算公式中 u 的选取，其中 u 是 L 上的非零向量. 为证明这一点，假定给定了 y 和 u，并且根据本节公式(2)已经计算出 $\hat{y}$. 在该公式中以 cu 代替 u，其中 c 是任意非零数量. 证明新公式的计算结果也是 $\hat{y}$.
32. 令 $\{v_1,v_2\}$ 是非零向量构成的正交集，c_1，c_2 是非零数量. 证明 $\{c_1v_1,c_2v_2\}$ 也是正交集. 由于集合的正交性由向量对来定义，这就表明如果将一个正交集中的向量施行标准化，得到的新集合仍然是正交的.
33. 给定 $\mathbf{R}^n$ 中的向量 $u\neq\mathbf{0}$，令 $L=\text{Span}\{u\}$. 证明映射 $x\mapsto \text{proj}_L x$ 是线性映射.
34. 给定 $\mathbf{R}^n$ 中的向量 $u\neq\mathbf{0}$，令 $L=\text{Span}\{u\}$. 对于 $y\in\mathbf{R}^n$，y 在 L 上的反射是点 $\text{refl}_L y$，其定义

为 $\text{refl}_L\boldsymbol{y}=2\cdot\text{proj}_L\boldsymbol{y}-\boldsymbol{y}$. 观察图像，可知 $\text{refl}_L\boldsymbol{y}$ 是 $\hat{\boldsymbol{y}}=\text{proj}_L\boldsymbol{y}$ 与 $\hat{\boldsymbol{y}}-\boldsymbol{y}$ 之和. 证明映射 $\boldsymbol{x}\mapsto\text{refl}_L\boldsymbol{x}$ 是线性映射.

35. [M] 证明：经过适当的矩阵变换可以使得矩阵 A 的列相互正交. 叙述你所用的变换.

$$A=\begin{bmatrix}-6&-3&6&1\\-1&2&1&-6\\3&6&3&-2\\6&-3&6&-1\\2&-1&2&3\\-3&6&3&2\\-2&-1&2&-3\\1&2&1&6\end{bmatrix}$$

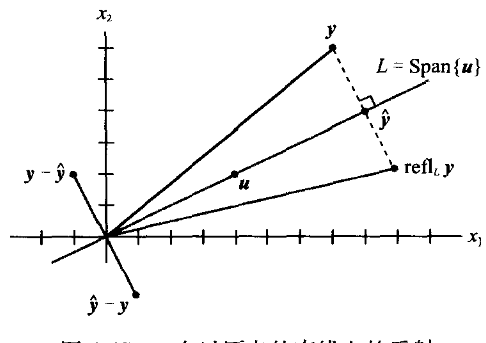


图6-17 $\boldsymbol{y}$ 在过原点的直线上的反射

36. [M] 在 a. ~d. 中，设 U 为对习题35矩阵 A 各列施行标准化后所得的矩阵.
   a. 计算 $U^{\text{T}}U$ 和 UU^{T}. 二者有何不同?
   b. 构造 $\mathbf{R}^8$ 的一个随机向量 $\boldsymbol{y}$，计算 $\boldsymbol{p}=UU^{\text{T}}\boldsymbol{y}$ 和 $\boldsymbol{z}=\boldsymbol{y}-\boldsymbol{p}$. 说明 $\boldsymbol{p}$ 为什么属于 Col A. 验证 $\boldsymbol{z}$ 与 $\boldsymbol{p}$ 正交.
   c. 验证 $\boldsymbol{z}$ 与 U 的每一列都正交.
   d. 注意到 $\boldsymbol{y}=\boldsymbol{p}+\boldsymbol{z}$ 且 $\boldsymbol{p}\in\text{Col }A$. 说明 $\boldsymbol{z}$ 为什么属于 $(\text{Col }A)^{\perp}$. （下一节中将说明 $\boldsymbol{y}$ 这一分解的重要性.）

基础练习答案

1. 向量是正交的，这是因为

$$\boldsymbol{u}_1\cdot\boldsymbol{u}_2=-2/5+2/5=0$$

它们是单位向量，因为

$$\|\boldsymbol{u}_1\|^2=(-1/\sqrt{5})^2+(2/\sqrt{5})^2=1/5+4/5=1$$

393

$$\|\boldsymbol{u}_2\|^2=(2/\sqrt{5})^2+(1/\sqrt{5})^2=4/5+1/5=1$$

特别地，集合 $\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$ 线性无关，又因为它包含两个向量，所以它是 $\mathbf{R}^2$ 的一组基.

2. 当 $\boldsymbol{y}=\begin{bmatrix}7\\6\end{bmatrix}$ 且 $\boldsymbol{u}=\begin{bmatrix}2\\1\end{bmatrix}$ 时，

$$\hat{\boldsymbol{y}}=\frac{\boldsymbol{y}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}}\boldsymbol{u}=\frac{20}{5}\begin{bmatrix}2\\1\end{bmatrix}=4\begin{bmatrix}2\\1\end{bmatrix}=\begin{bmatrix}8\\4\end{bmatrix}$$

这与例题3中求得的相同. 正交投影似乎不依赖于直线上 $\boldsymbol{u}$ 的选取，见习题31.

3. $$U\boldsymbol{y}=\begin{bmatrix}1/\sqrt{2}&2/3\\1/\sqrt{2}&-2/3\\0&1/3\end{bmatrix}\begin{bmatrix}-3\sqrt{2}\\6\end{bmatrix}=\begin{bmatrix}1\\-7\\2\end{bmatrix}$$

此外，由例题6可知，$\boldsymbol{x}=\begin{bmatrix}\sqrt{2}\\3\end{bmatrix}$ 及 $U\boldsymbol{x}=\begin{bmatrix}3\\-1\\1\end{bmatrix}$. 于是

$$U\boldsymbol{x}\cdot U\boldsymbol{y}=3+7+2=12\text{，且 }\boldsymbol{x}\cdot\boldsymbol{y}=-6+18=12$$

6.3　正 交 投 影

$\mathbf{R}^2$中一个点在过原点的直线上有正交投影，这一事实在$\mathbf{R}^n$上有重要的推广. 给定R^n中的向量$\boldsymbol{y}$和子空间W，存在W中的向量$\hat{\boldsymbol{y}}$使得：(1)$\hat{\boldsymbol{y}}$是W中使$\boldsymbol{y}-\hat{\boldsymbol{y}}$与$W$正交的唯一向量，且(2)$\hat{\boldsymbol{y}}$是$W$中与$\boldsymbol{y}$距离最近的唯一向量，如图6-18. $\hat{\boldsymbol{y}}$的这两个性质是求出本章实例介绍中线性方程组最小二乘解的关键. 详情将在6.5节阐明.

图　6-18

为给第一个定理做准备，我们观察当向量$\boldsymbol{y}$写成$\mathbf{R}^n$的一组基$\boldsymbol{u}_1$，…，$\boldsymbol{u}_n$的线性组合时，$\boldsymbol{y}$的和式中的项可以分成两部分，即，$\boldsymbol{y}$可以写成

$$\boldsymbol{y}=\boldsymbol{z}_1+\boldsymbol{z}_2$$

其中$\boldsymbol{z}_1$是某些$\boldsymbol{u}_i$的线性组合，$\boldsymbol{z}_2$则是其余$\boldsymbol{u}_i$的线性组合. 当$\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n\}$是正交基时这个想法尤其有用. 回忆6.1节中记号$W^{\perp}$，它表示与子空间$W$正交的全体向量构成的集合.

【例题1】　设$\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_5\}$是$\mathbf{R}^5$的一组正交基，且

$$\boldsymbol{y}=c_1\boldsymbol{u}_1+\cdots+c_5\boldsymbol{u}_5$$

考虑子空间$W=\mathrm{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$，将$\boldsymbol{y}$写成$W$中向量$\boldsymbol{z}_1$与$W^{\perp}$中向量$\boldsymbol{z}_2$之和.

394

解：记

$$\boldsymbol{y}=\underbrace{c_1\boldsymbol{u}_1+c_2\boldsymbol{u}_2}_{\boldsymbol{z}_1}+\underbrace{c_3\boldsymbol{u}_3+c_4\boldsymbol{u}_4+c_5\boldsymbol{u}_5}_{\boldsymbol{z}_2}$$

其中

$$\boldsymbol{z}_1=c_1\boldsymbol{u}_1+c_2\boldsymbol{u}_2\in\mathrm{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$$

且

$$\boldsymbol{z}_2=c_3\boldsymbol{u}_3+c_4\boldsymbol{u}_4+c_5\boldsymbol{u}_5\in\mathrm{Span}\{\boldsymbol{u}_3,\boldsymbol{u}_4,\boldsymbol{u}_5\}$$

为证$\boldsymbol{z}_2\in W^{\perp}$，只需证明$\boldsymbol{z}_2$与$W$的基$\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$中的向量正交(见6.1节). 根据内积的性质，因为$\boldsymbol{u}_1$正交于$\boldsymbol{u}_3$，$\boldsymbol{u}_4$和$\boldsymbol{u}_5$，可计算得：

$$\boldsymbol{z}_2\cdot\boldsymbol{u}_1=(c_3\boldsymbol{u}_3+c_4\boldsymbol{u}_4+c_5\boldsymbol{u}_5)\cdot\boldsymbol{u}_1=c_3\boldsymbol{u}_3\cdot\boldsymbol{u}_1+c_4\boldsymbol{u}_4\cdot\boldsymbol{u}_1+c_5\boldsymbol{u}_5\cdot\boldsymbol{u}_1=0$$

类似地可以证明$\boldsymbol{z}_2\cdot\boldsymbol{u}_2=0$. 因此$\boldsymbol{z}_2\in W^{\perp}$. ■

下面的定理表明，不事先给定$\mathbf{R}^n$的一组基仍然可以算出例题1中的分解$\boldsymbol{y}=\boldsymbol{z}_1+\boldsymbol{z}_2$. 只要给定$W$的一组正交基就足够了.

【定理8】　**正交分解定理**

设W是$\mathbf{R}^n$的一个子空间. 则$\mathbf{R}^n$的每一向量$\boldsymbol{y}$都能唯一写成

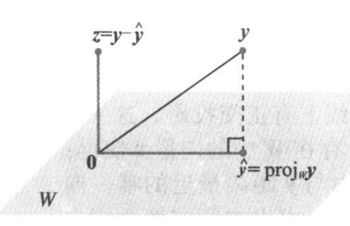
$$\boldsymbol{y}=\hat{\boldsymbol{y}}+\boldsymbol{z}\tag{1}$$

其中$\hat{\boldsymbol{y}}\in W$，$\boldsymbol{z}\in W^{\perp}$. 事实上，若$\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$是$W$的任意一组正交基，则

$$\hat{\boldsymbol{y}}=\frac{\boldsymbol{y}\cdot\boldsymbol{u}_1}{\boldsymbol{u}_1\cdot\boldsymbol{u}_1}\boldsymbol{u}_1+\cdots+\frac{\boldsymbol{y}\cdot\boldsymbol{u}_p}{\boldsymbol{u}_p\cdot\boldsymbol{u}_p}\boldsymbol{u}_p\tag{2}$$

并且$\boldsymbol{z}=\boldsymbol{y}-\hat{\boldsymbol{y}}$.

(1)中的向量$\hat{\boldsymbol{y}}$被称作**$\boldsymbol{y}$在W上的正交投影**(orthogonal projection of $\boldsymbol{y}$ onto W)，通常记作$\mathrm{proj}_W\boldsymbol{y}$，如图6-19. 当$W$是1维子空间时，$\hat{\boldsymbol{y}}$的计算公式与6.2节所给公式一致.



395

图6-19　y在W上的正交投影

证明：设$\{u_1,\cdots,u_p\}$是W的一组正交基，定义$\hat{y}$如(2)式.[1] 由于$\hat{y}$是基的线性组合，所以$\hat{y}\in W$. 令$z=y-\hat{y}$. 因为u_1与u_2，…，u_p正交，由(2)可知：

$$z\cdot u_1=(y-\hat{y})\cdot u_1=y\cdot u_1-\left(\frac{y\cdot u_1}{u_1\cdot u_1}\right)u_1\cdot u_1-0-\cdots-0$$
$$=y\cdot u_1-y\cdot u_1=0$$

于是z与u_1正交. 类似地，z与W的一组基中的每个向量u_j正交. 因此z与W中任意向量正交，即$z\in W^{\perp}$.

为证明(1)中分解唯一，设y可以写成$y=\hat{y}_1+z_1$，其中$\hat{y}_1\in W$，$z_1\in W^{\perp}$. 则$\hat{y}+z=\hat{y}_1+z_1$(因为两端都等于y)，于是

$$\hat{y}-\hat{y}_1=z_1-z$$

这个等式表明向量$v=\hat{y}-\hat{y}_1$属于W，也属于$W^{\perp}$(因为z_1和z都属于$W^{\perp}$，而$W^{\perp}$是子空间). 所以$v\cdot v=0$，于是$v=\mathbf{0}$. 这就证明了$\hat{y}=\hat{y}_1$并且$z_1=z$. ■

分解(1)的唯一性表明，正交投影$\hat{y}$只依赖于W，不依赖于(2)中某组特定的基.

【例题2】　令$u_1=\begin{bmatrix}2\\5\\-1\end{bmatrix}$，$u_2=\begin{bmatrix}-2\\1\\1\end{bmatrix}$以及$y=\begin{bmatrix}1\\2\\3\end{bmatrix}$. 注意到$\{u_1,u_2\}$是$W=\mathrm{Span}\{u_1,u_2\}$的一组正交基. 将$y$写成两向量之和，其中一个向量属于$W$，一个向量正交于$W$.

解：y在W上的正交投影是：

$$\hat{y}=\frac{y\cdot u_1}{u_1\cdot u_1}u_1+\frac{y\cdot u_2}{u_2\cdot u_2}u_2=\frac{9}{30}\begin{bmatrix}2\\5\\-1\end{bmatrix}+\frac{3}{6}\begin{bmatrix}-2\\1\\1\end{bmatrix}=\frac{9}{30}\begin{bmatrix}2\\5\\-1\end{bmatrix}+\frac{15}{30}\begin{bmatrix}-2\\1\\1\end{bmatrix}=\begin{bmatrix}-2/5\\2\\1/5\end{bmatrix}$$

还有

$$y-\hat{y}=\begin{bmatrix}1\\2\\3\end{bmatrix}-\begin{bmatrix}-2/5\\2\\1/5\end{bmatrix}=\begin{bmatrix}7/5\\0\\14/5\end{bmatrix}$$

定理8已经保证$y-\hat{y}\in W^{\perp}$. 不过若要加以验证，一个很好的方法是验证$y-\hat{y}$与u_1

---

1. 我们可不妨假定W不是零空间，否则，$W^{\perp}=\mathbf{R}^n$并且(1)式就是$y=\mathbf{0}+y$. 下节将证明：$\mathbf{R}^n$的任意非零子空间都有一组正交基.

及 $\boldsymbol{u}_2$ 都正交，从而与整个 W 都正交. 所求 $\boldsymbol{y}$ 的分解为：

$$\boldsymbol{y}=\begin{bmatrix}1\\2\\3\end{bmatrix}=\begin{bmatrix}-2/5\\2\\1/5\end{bmatrix}+\begin{bmatrix}7/5\\0\\14/5\end{bmatrix}$$ ■

396

6.3.1 正交投影的几何解释

当 W 是 1 维子空间时，$\text{proj}_W\,\boldsymbol{y}$ 的计算公式(2)仅包含一项. 这样，当 $\dim W>1$ 时，(2)式中的每一项本身都是 $\boldsymbol{y}$ 在某个 1 维子空间上的正交投影，这个子空间由 W 基底中的某个 $\boldsymbol{u}$ 张成. 图 6-20 描述了当 W 是 $\mathbf{R}^3$ 中由 $\boldsymbol{u}_1$，$\boldsymbol{u}_2$ 张成的子空间时的情形. 这里 $\hat{\boldsymbol{y}}_1$ 和 $\hat{\boldsymbol{y}}_2$ 分别表示 $\boldsymbol{y}$ 在由 $\boldsymbol{u}_1$ 和 $\boldsymbol{u}_2$ 张成的直线上的投影. $\boldsymbol{y}$ 在 W 上的正交投影 $\hat{\boldsymbol{y}}$ 就是 $\boldsymbol{y}$ 在这两个彼此正交的 1 维子空间上的投影之和.

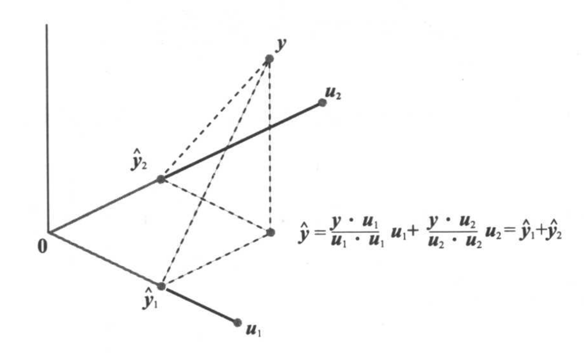


图 6-20 $\boldsymbol{y}$ 的正交投影等于它在彼此正交的 1 维子空间上的投影之和

6.3.2 正交投影的性质

若 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是 W 的一组正交基，且 $\boldsymbol{y}$ 恰好属于 W，则 $\text{proj}_W\,\boldsymbol{y}$ 的计算公式恰与 6.2 节定理 5 所给的 $\boldsymbol{y}$ 的表达式一致. 此时有 $\text{proj}_W\,\boldsymbol{y}=\boldsymbol{y}$.

若 $\boldsymbol{y}\in W=\text{Span}\{\boldsymbol{u}_1\cdots,\boldsymbol{u}_p\}$，则 $\text{proj}_W\,\boldsymbol{y}=\boldsymbol{y}$.

这一事实也可由下面的定理推知.

397

【定理 9】 最佳逼近定理

设 W 是 $\mathbf{R}^n$ 的一个子空间，$\boldsymbol{y}$ 是 $\mathbf{R}^n$ 中任意向量，$\hat{\boldsymbol{y}}$ 是 $\boldsymbol{y}$ 在 W 上的正交投影. 则 $\hat{\boldsymbol{y}}$ 是 W 中与 $\boldsymbol{y}$ 距离最近的点，即，对 W 中任何异于 $\hat{\boldsymbol{y}}$ 的向量 $\boldsymbol{v}$，都有：

$$\|\boldsymbol{y}-\hat{\boldsymbol{y}}\|<\|\boldsymbol{y}-\boldsymbol{v}\| \tag{3}$$

定理 9 中的向量 $\hat{\boldsymbol{y}}$ 被称作 **W 中元素对 $\boldsymbol{y}$ 的最佳逼近**(the best approximation to $\boldsymbol{y}$ by elements of W). 在后续章节中我们将要探讨这样一类问题，其中，一个已知向量 $\boldsymbol{y}$ 必须用固定子空间 W 中的某个向量 $\boldsymbol{v}$ 来替代或逼近. $\boldsymbol{y}$ 到 $\boldsymbol{v}$ 的距离由 $\|\boldsymbol{y}-\boldsymbol{v}\|$ 给出，它可以看作是用 $\boldsymbol{v}$ 替代 $\boldsymbol{y}$ 的“误差”. 定理 9 表明，当 $\boldsymbol{v}=\hat{\boldsymbol{y}}$ 时误差达到最小值.

(3)式给出了一种新的方法，可以证明 $\hat{\boldsymbol{y}}$ 不依赖于计算时选取的特定正交基. 如果另选 W 的一组正交基来计算 $\boldsymbol{y}$ 的正交投影，那么这个投影应该也是 W 中与 $\boldsymbol{y}$ 距离最近的点，也就是 $\hat{\boldsymbol{y}}$.

证明：选取 W 中异于 $\hat{\boldsymbol{y}}$ 的向量 $\boldsymbol{v}$，如图6-21．则 $\hat{\boldsymbol{y}}-\boldsymbol{v}\in W$．根据正交分解定理可知，$\boldsymbol{y}-\hat{\boldsymbol{y}}$与 W 正交．特别地，$\boldsymbol{y}-\hat{\boldsymbol{y}}$ 与 $\hat{\boldsymbol{y}}-\boldsymbol{v}$(属于 W)正交．因为

$$\boldsymbol{y}-\boldsymbol{v}=(\boldsymbol{y}-\hat{\boldsymbol{y}})+(\hat{\boldsymbol{y}}-\boldsymbol{v})$$

由勾股定理有

$$\|\boldsymbol{y}-\boldsymbol{v}\|^2=\|\boldsymbol{y}-\hat{\boldsymbol{y}}\|^2+\|\hat{\boldsymbol{y}}-\boldsymbol{v}\|^2$$

(见图6-21中右侧的三角形，其各边的长度都已标出.)这样，由 $\hat{\boldsymbol{y}}-\boldsymbol{v}\neq 0$ 就有 $\|\hat{\boldsymbol{y}}-\boldsymbol{v}\|^2>0$，于是(3)中不等式成立．■

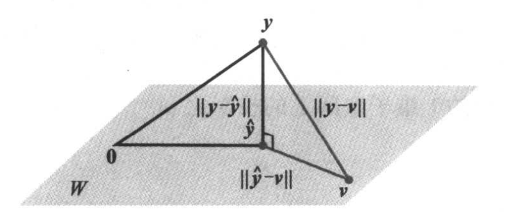


图6-21　$\boldsymbol{y}$ 在 W 上的正交投影是 W 中与 $\boldsymbol{y}$ 距离最近的点

【例题3】　与例题2相同，令 $\boldsymbol{u}_1=\begin{bmatrix}2\\5\\-1\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}-2\\1\\1\end{bmatrix}$，$\boldsymbol{y}=\begin{bmatrix}1\\2\\3\end{bmatrix}$以及 $W=\text{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$，则 W 中与 $\boldsymbol{y}$ 距离最近的向量是：

398

$$\hat{\boldsymbol{y}}=\frac{\boldsymbol{y}\cdot\boldsymbol{u}_1}{\boldsymbol{u}_1\cdot\boldsymbol{u}_1}\boldsymbol{u}_1+\frac{\boldsymbol{y}\cdot\boldsymbol{u}_2}{\boldsymbol{u}_2\cdot\boldsymbol{u}_2}\boldsymbol{u}_2=\begin{bmatrix}-2/5\\2\\1/5\end{bmatrix}$$ ■

【例题4】　$\mathbb{R}^n$中一点 $\boldsymbol{y}$ 到子空间 W 的距离被定义成 $\boldsymbol{y}$ 到 W 中最近一点的距离．试求 $\boldsymbol{y}$ 到 $W=\text{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$的距离，其中

$$\boldsymbol{y}=\begin{bmatrix}-1\\-5\\10\end{bmatrix},\quad \boldsymbol{u}_1=\begin{bmatrix}5\\-2\\1\end{bmatrix},\quad \boldsymbol{u}_2=\begin{bmatrix}1\\2\\-1\end{bmatrix}$$

解：由最佳逼近定理可知，$\boldsymbol{y}$ 到 W 的距离是$\|\boldsymbol{y}-\hat{\boldsymbol{y}}\|$，其中 $\hat{\boldsymbol{y}}=\text{proj}_W\boldsymbol{y}$．由于$\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$是 W 的一组正交基，因此有：

$$\hat{\boldsymbol{y}}=\frac{15}{30}\boldsymbol{u}_1+\frac{-21}{6}\boldsymbol{u}_2=\frac{1}{2}\begin{bmatrix}5\\-2\\1\end{bmatrix}-\frac{7}{2}\begin{bmatrix}1\\2\\-1\end{bmatrix}=\begin{bmatrix}-1\\-8\\4\end{bmatrix}$$

$$\boldsymbol{y}-\hat{\boldsymbol{y}}=\begin{bmatrix}-1\\-5\\10\end{bmatrix}-\begin{bmatrix}-1\\-8\\4\end{bmatrix}=\begin{bmatrix}0\\3\\6\end{bmatrix}$$

$$\|\boldsymbol{y}-\hat{\boldsymbol{y}}\|^2=3^2+6^2=45$$

所以 $\boldsymbol{y}$ 到 W 的距离是 $\sqrt{45}=3\sqrt{5}$．■

本节最后一个定理将要揭示：当选取的基是 W 的标准正交集时，$\text{proj}_W\,\boldsymbol{y}$ 的计算公式(2)可以如何简化.

【定理 10】 若$\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$是 $\mathbf{R}^n$子空间 W 的一组标准正交基，则

$$\text{proj}_W\boldsymbol{y}=(\boldsymbol{y}\cdot\boldsymbol{u}_1)\boldsymbol{u}_1+(\boldsymbol{y}\cdot\boldsymbol{u}_2)\boldsymbol{u}_2+\cdots+(\boldsymbol{y}\cdot\boldsymbol{u}_p)\boldsymbol{u}_p \tag{4}$$

若 $U=[\boldsymbol{u}_1\quad \boldsymbol{u}_2\quad\cdots\quad \boldsymbol{u}_p]$，则

$$\text{对任意 } \boldsymbol{y}\in\mathbf{R}^n，\text{有 } \text{proj}_W\boldsymbol{y}=UU^{\mathrm{T}}\boldsymbol{y} \tag{5}$$

证明： 公式(4)可直接由(2)推得. 此外，(4)式表明 $\text{proj}_W\,\boldsymbol{y}$ 是 U 中列向量的线性组合，其权重为 $\boldsymbol{y}\cdot\boldsymbol{u}_1$，$\boldsymbol{y}\cdot\boldsymbol{u}_2$，…，$\boldsymbol{y}\cdot\boldsymbol{u}_p$. 它们可以写作 $\boldsymbol{u}_1^{\mathrm{T}}\cdot\boldsymbol{y}$，$\boldsymbol{u}_2^{\mathrm{T}}\cdot\boldsymbol{y}$，…，$\boldsymbol{u}_p^{\mathrm{T}}\cdot\boldsymbol{y}$，因而是 $U^{\mathrm{T}}\boldsymbol{y}$ 中的元素，这就证明了公式(5). ■

设 U 是一个有标准正交列的 $n\times p$ 矩阵，令 W 是 U 的列空间. 则

$$\text{对任意 } \boldsymbol{x}\in\mathbf{R}^p，\text{有 } UU^{\mathrm{T}}\boldsymbol{x}=I_p\boldsymbol{x}=\boldsymbol{x}\qquad\text{定理 6}$$

$$\text{对任意 } \boldsymbol{y}\in\mathbf{R}^n，\text{有 } UU^{\mathrm{T}}\boldsymbol{y}=\text{proj}_W\,\boldsymbol{y}\qquad\text{定理 10}$$

若 U 是一个有标准正交列的 $n\times n$(方)阵，则 U 是正交矩阵，其列空间 W 是整个 $\mathbf{R}^n$，并且对任意 $\boldsymbol{y}\in\mathbf{R}^n$，有 $UU^{\mathrm{T}}\boldsymbol{y}=I\boldsymbol{y}=\boldsymbol{y}$.

尽管从理论意义上来说公式(4)相当重要，但在实际应用中它常常涉及数($\boldsymbol{u}_i$中的元素)的平方根运算. 因此手算时推荐使用公式(2).

399

基础练习

令$\boldsymbol{u}_1=\begin{bmatrix}-7\\1\\4\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}-1\\1\\-2\end{bmatrix}$，$\boldsymbol{y}=\begin{bmatrix}-9\\1\\6\end{bmatrix}$以及 $W=\text{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$. 利用 $\boldsymbol{u}_1$与 $\boldsymbol{u}_2$正交的事实，计算 $\text{proj}_W\,\boldsymbol{y}$.

习题 6.3

在习题 1 和 2 中，你可以假定$\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_4\}$是 $\mathbf{R}^4$的一组正交基.

1. $\boldsymbol{u}_1=\begin{bmatrix}0\\1\\-4\\-1\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}3\\5\\1\\1\end{bmatrix}$，$\boldsymbol{u}_3=\begin{bmatrix}1\\0\\1\\-4\end{bmatrix}$，$\boldsymbol{u}_4\begin{bmatrix}5\\-3\\-1\\1\end{bmatrix}$，$\boldsymbol{x}=\begin{bmatrix}10\\-8\\2\\0\end{bmatrix}$. 将 $\boldsymbol{x}$ 写成两向量之和，其中一个属于 $\text{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3\}$，另一个属于 $\text{Span}\{\boldsymbol{u}_4\}$.

2. $\boldsymbol{u}_1=\begin{bmatrix}1\\2\\1\\1\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}-2\\1\\-1\\1\end{bmatrix}$，$\boldsymbol{u}_3=\begin{bmatrix}1\\1\\-2\\-1\end{bmatrix}$，$\boldsymbol{u}_4=\begin{bmatrix}-1\\1\\1\\-2\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}4\\5\\-3\\3\end{bmatrix}$. 将 $\boldsymbol{v}$ 写成两向量之和，其中一个属于 $\text{Span}\{\boldsymbol{u}_1\}$，另一个属于 $\text{Span}\{\boldsymbol{u}_2,\boldsymbol{u}_3,\boldsymbol{u}_4\}$.

验证习题 3 ~ 6 中的$\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$是正交集，然后求 $\boldsymbol{y}$ 在 $\text{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$上的正交投影.

3. $\boldsymbol{y}=\begin{bmatrix}-1\\4\\3\end{bmatrix}$，$\boldsymbol{u}_1=\begin{bmatrix}1\\1\\0\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}-1\\1\\0\end{bmatrix}$

4. $\boldsymbol{y}=\begin{bmatrix}6\\3\\-2\end{bmatrix}$，$\boldsymbol{u}_1=\begin{bmatrix}3\\4\\0\end{bmatrix}$，$\boldsymbol{u}_2=\begin{bmatrix}-4\\3\\0\end{bmatrix}$

5. $\boldsymbol{y}=\begin{bmatrix}-1\\2\\6\end{bmatrix}$, $\boldsymbol{u}_1=\begin{bmatrix}3\\-1\\2\end{bmatrix}$, $\boldsymbol{u}_2=\begin{bmatrix}1\\-1\\-2\end{bmatrix}$

6. $\boldsymbol{y}=\begin{bmatrix}6\\4\\1\end{bmatrix}$, $\boldsymbol{u}_1=\begin{bmatrix}-4\\-1\\1\end{bmatrix}$, $\boldsymbol{u}_2=\begin{bmatrix}0\\1\\1\end{bmatrix}$

在习题7~10中，令 W 是由 $\boldsymbol{u}$ 张成的子空间. 将 $\boldsymbol{y}$ 写成两向量之和，其中一个属于 W，另一个与 W 正交.

7. $\boldsymbol{y}=\begin{bmatrix}1\\3\\5\end{bmatrix}$, $\boldsymbol{u}_1=\begin{bmatrix}1\\3\\-2\end{bmatrix}$, $\boldsymbol{u}_2=\begin{bmatrix}5\\1\\4\end{bmatrix}$

8. $\boldsymbol{y}=\begin{bmatrix}-1\\4\\3\end{bmatrix}$, $\boldsymbol{u}_1=\begin{bmatrix}1\\1\\1\end{bmatrix}$, $\boldsymbol{u}_2=\begin{bmatrix}-1\\3\\-2\end{bmatrix}$

9. $\boldsymbol{y}=\begin{bmatrix}4\\3\\3\\-1\end{bmatrix}$, $\boldsymbol{u}_1=\begin{bmatrix}1\\1\\0\\1\end{bmatrix}$, $\boldsymbol{u}_2=\begin{bmatrix}-1\\3\\1\\-2\end{bmatrix}$, $\boldsymbol{u}_3=\begin{bmatrix}-1\\0\\1\\1\end{bmatrix}$

10. $\boldsymbol{y}=\begin{bmatrix}3\\4\\5\\6\end{bmatrix}$, $\boldsymbol{u}_1=\begin{bmatrix}1\\1\\0\\-1\end{bmatrix}$, $\boldsymbol{u}_2=\begin{bmatrix}1\\0\\1\\1\end{bmatrix}$, $\boldsymbol{u}_3=\begin{bmatrix}0\\-1\\1\\-1\end{bmatrix}$

在习题11和12中，求 $\boldsymbol{v}_1$，$\boldsymbol{v}_2$所张成的子空间 W 中，与 $\boldsymbol{y}$ 距离最近的点.

11. $\boldsymbol{y}=\begin{bmatrix}3\\1\\5\\1\end{bmatrix}$, $\boldsymbol{v}_1=\begin{bmatrix}3\\1\\-1\\1\end{bmatrix}$, $\boldsymbol{v}_2=\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}$

12. $\boldsymbol{y}=\begin{bmatrix}3\\-1\\1\\13\end{bmatrix}$, $\boldsymbol{v}_1=\begin{bmatrix}1\\-2\\-1\\2\end{bmatrix}$, $\boldsymbol{v}_2=\begin{bmatrix}-4\\1\\0\\3\end{bmatrix}$

400

在习题13和14中，求 $c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2$形式的向量对 $\boldsymbol{y}$ 的最佳逼近.

13. $\boldsymbol{z}=\begin{bmatrix}3\\-7\\2\\3\end{bmatrix}$, $\boldsymbol{v}_1=\begin{bmatrix}2\\-1\\-3\\1\end{bmatrix}$, $\boldsymbol{v}_2=\begin{bmatrix}1\\1\\0\\-1\end{bmatrix}$

14. $\boldsymbol{z}=\begin{bmatrix}2\\4\\0\\-1\end{bmatrix}$, $\boldsymbol{v}_1=\begin{bmatrix}2\\0\\-1\\-3\end{bmatrix}$, $\boldsymbol{v}_2=\begin{bmatrix}5\\-2\\4\\2\end{bmatrix}$

15. 令 $\boldsymbol{y}=\begin{bmatrix}5\\-9\\5\end{bmatrix}$, $\boldsymbol{u}_1=\begin{bmatrix}-3\\-5\\1\end{bmatrix}$, $\boldsymbol{u}_2=\begin{bmatrix}-3\\2\\1\end{bmatrix}$. 求 $\boldsymbol{y}$ 到 $\mathbf{R}^3$ 中 $\boldsymbol{u}_1$，$\boldsymbol{u}_2$所张成平面的距离.

16. 设 $\boldsymbol{y}$，$\boldsymbol{v}_1$和 $\boldsymbol{v}_2$同习题12. 求 $\boldsymbol{y}$ 到 $\mathbf{R}^4$ 中 $\boldsymbol{v}_1$，$\boldsymbol{v}_2$所张成子空间的距离.

17. 令 $\boldsymbol{y}=\begin{bmatrix}4\\8\\1\end{bmatrix}$, $\boldsymbol{u}_1=\begin{bmatrix}2/3\\1/3\\2/3\end{bmatrix}$, $\boldsymbol{u}_2=\begin{bmatrix}-2/3\\2/3\\1/3\end{bmatrix}$且 $W=\mathrm{Span}\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$.

a. 设 $U=[\boldsymbol{u}_1\quad \boldsymbol{u}_2]$，计算 $U^{\mathrm{T}}U$ 和 UU^{T}.

b. 计算 $\mathrm{proj}_W\,\boldsymbol{y}$ 和$(UU^{\mathrm{T}})\boldsymbol{y}$.

18. 令 $\boldsymbol{y}=\begin{bmatrix}7\\9\end{bmatrix}$, $\boldsymbol{u}_1=\begin{bmatrix}1/\sqrt{10}\\-3/\sqrt{10}\end{bmatrix}$且 $W=\mathrm{Span}\{\boldsymbol{u}_1\}$.

a. 设 U 是只包含列 u_1 的 2×1 矩阵，计算 U^TU 和 UU^T.

b. 计算 $\text{proj}_W y$ 和$(UU^T)y$.

19. 令 $u_1=\begin{bmatrix}1\\1\\-2\end{bmatrix}$，$u_2=\begin{bmatrix}5\\-1\\2\end{bmatrix}$以及 $u_3=\begin{bmatrix}0\\0\\1\end{bmatrix}$. 注意 u_1与 u_2正交，但 u_3与 u_1或者 u_2都不正交. 可以证明 u_3不属于 u_1和 u_2所张成的子空间 W. 利用这个事实构造 $\mathbf{R}^3$的一个非零向量 v，使之与 u_1，u_2都正交.

20. 设 u_1和 u_2同习题 19，且 $u_4=\begin{bmatrix}0\\1\\0\end{bmatrix}$. 可以证明 u_4不属于 u_1和 u_2所张成的子空间 W. 利用这个事实构造 $\mathbf{R}^3$的一个非零向量 v，使之与 u_1，u_2都正交.

习题 21 和 22 中的所有向量和子空间都在 $\mathbf{R}^n$中. 判断各命题的真假，并说明理由.

21. a. 若 z 与 u_1和 u_2正交，且 $W=\text{Span}\{u_1,u_2\}$，则 z 一定属于 $W^\perp$.
    b. 对每个 y 及子空间 W，向量 $y\text{-proj}_W y$ 正交于 W.
    c. y 在子空间W上的正交投影 $\hat{y}$ 有时依赖于计算 $\hat{y}$ 时用到的 W 的基.
    d. 若 y 属于子空间 W，则 y 在 W 上的正交投影就是它自身.
    e. 若 $n\times p$ 矩阵 U 的各列标准正交，则 UU^Ty 是 y 在 U 的列空间上的正交投影.

22. a. 若 W 是 $\mathbf{R}^n$的子空间且 v 同时属于 W 和 $W^\perp$，则 v 一定是零向量.
    b. 正交分解定理中，$\hat{y}$ 的计算公式(2)的每一项，其本身都是 y 在 W 的一个子空间上的正交投影.
    c. 若 $y=z_1+z_2$，其中 $z_1\in W$，$z_2\in W^\perp$，则 z_1就是 y 在 W 上的正交投影.
    d. W 中元素对 y 的最佳逼近是向量 $y\text{-proj}_W y$.
    e. 若 $n\times p$ 矩阵 U 有标准正交列，则对任意 $x\in\mathbf{R}^n$，有 $UU^Tx=x$.

23. 设 A 是 $m\times n$ 矩阵. 证明 $\mathbf{R}^n$中任意向量 x 可以写成 $x=p+u$ 的形式，其中 $p\in\text{Row }A$，$u\in\text{Nul }A$. 此外，证明如果方程 $Ax=b$ 相容，则 Row A 中存在唯一的 p 使 $Ap=b$.

24. 设 W 是以$\{w_1,\cdots,w_p\}$为正交基的 $\mathbf{R}^n$的子空间，又设$\{v_1,\cdots,v_q\}$是 $W^\perp$ 的一组正交基.
    a. 说明$\{w_1,\cdots,w_p,v_1,\cdots,v_q\}$是正交集.
    b. 说明(a)中的集合张成 $\mathbf{R}^n$.
    c. 证明 $\dim W+\dim W^\perp=n$.

25. [M]设 U 为 6.2 节习题 36 中的 8×4 矩阵. 求 Col U 中与 $y=(1,1,1,1,1,1,1,1)$距离最近的点. 写出求解该问题时你所用的按键和命令.

26. [M]设 U 为习题 25 中的矩阵，求 $b=(1,1,1,1,-1,-1,-1,-1)$到 Col U 的距离. 401

基础练习答案

计算

$$\text{proj}_W y=\frac{y\cdot u_1}{u_1\cdot u_1}u_1+\frac{y\cdot u_2}{u_2\cdot u_2}u_2=\frac{88}{66}u_1+\frac{-2}{6}u_2$$

$$=\frac{4}{3}\begin{bmatrix}-7\\1\\4\end{bmatrix}-\frac{1}{3}\begin{bmatrix}-1\\1\\-2\end{bmatrix}=\begin{bmatrix}-9\\1\\6\end{bmatrix}=y$$

这里 y 恰好是 u_1和 u_2的一个线性组合，因此 $y\in W$. W 中与 y 距离最近的点就是 y 本身.

6.4　格拉姆－施密特方法

格拉姆－施密特方法是构造$\mathbf{R}^n$中任意非零子空间的正交基或标准正交基的一种简单算法．本节前两道例题演示了运用这一方法的手算过程．

【例题 1】　令 $W=\text{Span}\{\boldsymbol{x}_1,\boldsymbol{x}_2\}$，其中 $\boldsymbol{x}_1=\begin{bmatrix}3\\6\\0\end{bmatrix}$且$\boldsymbol{x}_2=\begin{bmatrix}1\\2\\2\end{bmatrix}$．构造 W 的一组正交基 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$．

解： 子空间 W，向量 $\boldsymbol{x}_1$、$\boldsymbol{x}_2$及 $\boldsymbol{x}_2$在 $\boldsymbol{x}_1$上的投影 $\boldsymbol{p}$ 如图 6-22 所示．$\boldsymbol{x}_2$中正交于 $\boldsymbol{x}_1$的分量是$\boldsymbol{x}_2-\boldsymbol{p}$，因为$\boldsymbol{x}_2-\boldsymbol{p}$ 由 $\boldsymbol{x}_2$和 $\boldsymbol{x}_1$的一个数乘构成，它仍然属于 W．令 $\boldsymbol{v}_1=\boldsymbol{x}_1$且

$$\boldsymbol{v}_2=\boldsymbol{x}_2-\boldsymbol{p}=\boldsymbol{x}_2-\frac{\boldsymbol{x}_2\cdot\boldsymbol{x}_1}{\boldsymbol{x}_1\cdot\boldsymbol{x}_1}\boldsymbol{x}_1=\begin{bmatrix}1\\2\\2\end{bmatrix}-\frac{15}{45}\begin{bmatrix}3\\6\\0\end{bmatrix}=\begin{bmatrix}0\\0\\2\end{bmatrix}$$

则$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$是 W 中非零向量构成的正交集．由于 $\dim W=2$，集合$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$是 W 的一组基．■

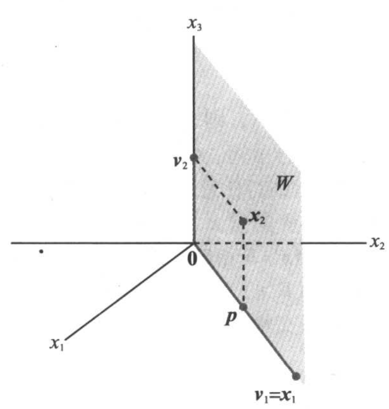


图 6-22　构造 W 的一组基$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$

下面的例题演示了格拉姆－施密特方法的全过程，请认真学习．

【例题 2】　令 $\boldsymbol{x}_1=\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$，$\boldsymbol{x}_2=\begin{bmatrix}0\\1\\1\\1\end{bmatrix}$，$\boldsymbol{x}_3=\begin{bmatrix}0\\0\\1\\1\end{bmatrix}$．则$\{\boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \boldsymbol{x}_3\}$显然线性无关，并且是$\mathbf{R}^4$的某个子空间 W 的基．试构造 W 的一组正交基．

解：

402

步骤 1. 令 $\boldsymbol{v}_1=\boldsymbol{x}_1$，且 $W_1=\text{Span}\{\boldsymbol{x}_1\}=\text{Span}\{\boldsymbol{v}_1\}$．

步骤 2. 令 $\boldsymbol{v}_2$是 $\boldsymbol{x}_2$减去 $\boldsymbol{x}_2$在子空间 W_1上的投影所得的向量．即，令

$$\begin{aligned}\boldsymbol{v}_2 &= \boldsymbol{x}_2 - \text{proj}_{W_1}\boldsymbol{x}_2 \\ &= \boldsymbol{x}_2 - \frac{\boldsymbol{x}_2 \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 \quad \text{由于 } \boldsymbol{v}_1 = \boldsymbol{x}_1 \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}\end{aligned}$$

与例题 1 相同，$\boldsymbol{v}_2$是$\boldsymbol{x}_2$中正交于$\boldsymbol{x}_1$的分量，且$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$是$\boldsymbol{x}_1$，$\boldsymbol{x}_2$所张成子空间 W_2的正交基.

步骤 2′(可选). 适当缩放 $\boldsymbol{v}_2$，以便简化后续计算. 因为 $\boldsymbol{v}_2$ 的元素包含分数，用 4 作为因子缩放 $\boldsymbol{v}_2$ 得 $\boldsymbol{v}'_2$，并用正交基$\{\boldsymbol{v}_1,\boldsymbol{v}'_2\}$替换基$\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$.

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{v}'_2 = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

步骤 3. 令 $\boldsymbol{v}_3$是$\boldsymbol{x}_3$减去$\boldsymbol{x}_3$在子空间 W_2上的投影所得的向量. 利用正交基$\{\boldsymbol{v}_1,\boldsymbol{v}'_2\}$计算$\boldsymbol{x}_3$在 W_2上的投影：

$$\begin{aligned}\text{proj}_{W_2}\boldsymbol{x}_3 &= \underbrace{\frac{\boldsymbol{x}_3 \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1}_{\boldsymbol{x}_3\text{在 }\boldsymbol{v}_1\text{上的投影}} + \underbrace{\frac{\boldsymbol{x}_3 \cdot \boldsymbol{v}'_2}{\boldsymbol{v}'_2 \cdot \boldsymbol{v}'_2}\boldsymbol{v}'_2}_{\boldsymbol{x}_3\text{在 }\boldsymbol{v}'_2\text{ 上的投影}} \\ &= \frac{2}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{2}{12}\begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix}\end{aligned}$$

则 $\boldsymbol{v}_3$是$\boldsymbol{x}_3$中正交于 W_2的分量，即

$$\boldsymbol{v}_3 = \boldsymbol{x}_3 - \text{proj}_{W_2}\boldsymbol{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

这一构造过程如图 6-23 所示. 观察可知 $\boldsymbol{v}_3 \in W$，这是因为$\boldsymbol{x}_3$和 $\text{proj}_{W_2}\boldsymbol{x}_3$ 都属于 W. 这样$\{\boldsymbol{v}_1,\boldsymbol{v}'_2,\boldsymbol{v}_3\}$是非零向量构成的正交集，因此也是 W 中的线性无关集. 注意到 W 由三向量组成的基底所定义，于是 W维数为3. 因此，根据4.5节的基本定

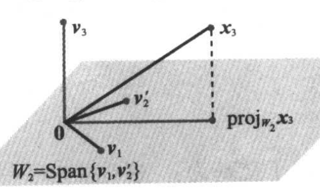


图 6-23 用$\boldsymbol{x}_3$和 W_2构造 $\boldsymbol{v}_3$

403 理可知，$\{\boldsymbol{v}_1, \boldsymbol{v}'_2, \boldsymbol{v}_3\}$是 W 的一组正交基. ■

下面定理的证明表明格拉姆－施密特方法是可行的. 由于缩放向量只是为了简化手算，因此下面略去这一步骤.

【定理 11】 格拉姆－施密特方法

给定 $\mathbf{R}^n$子空间 W 的一组基$\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_p\}$，定义：

$$\boldsymbol{v}_1 = \boldsymbol{x}_1$$

$$\boldsymbol{v}_2 = \boldsymbol{x}_2 - \frac{\boldsymbol{x}_2 \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1$$

$$\boldsymbol{v}_3 = \boldsymbol{x}_3 - \frac{\boldsymbol{x}_3 \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 - \frac{\boldsymbol{x}_3 \cdot \boldsymbol{v}_2}{\boldsymbol{v}_2 \cdot \boldsymbol{v}_2}\boldsymbol{v}_2$$

$$\vdots$$

$$\boldsymbol{v}_p = \boldsymbol{x}_p - \frac{\boldsymbol{x}_p \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 - \frac{\boldsymbol{x}_p \cdot \boldsymbol{v}_2}{\boldsymbol{v}_2 \cdot \boldsymbol{v}_2}\boldsymbol{v}_2 - \cdots - \frac{\boldsymbol{x}_p \cdot \boldsymbol{v}_{p-1}}{\boldsymbol{v}_{p-1} \cdot \boldsymbol{v}_{p-1}}\boldsymbol{v}_{p-1}$$

则$\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p\}$是 W 的一组正交基. 并且还有

$$\text{Span}\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\} = \text{Span}\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k\}, \quad 对\ 1 \leqslant k \leqslant p \tag{1}$$

证明： 对 $1 \leqslant k \leqslant p$，令 $W_k = \text{Span}\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k\}$. 设 $\boldsymbol{v}_1 = \boldsymbol{x}_1$，于是 $\text{Span}\{\boldsymbol{v}_1\} = \text{Span}\{\boldsymbol{x}_1\}$. 假定对某个 $k < p$，我们已经构造了 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k$使得$\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\}$构成 W_k的一组正交基. 定义：

$$\boldsymbol{v}_{k+1} = \boldsymbol{x}_{k+1} - \text{proj}_{W_k}\boldsymbol{x}_{k+1} \tag{2}$$

根据正交分解定理，$\boldsymbol{v}_{k+1}$与 W_k正交. 注意到 $\text{proj}_{W_k}\boldsymbol{x}_{k+1}$属于 W_k，从而属于 W_{k+1}. 又由于 $\boldsymbol{x}_{k+1} \in W_{k+1}$，所以 $\boldsymbol{v}_{k+1} \in W_{k+1}$(因为 W_{k+1}是子空间，对减法封闭). 此外，由 $\boldsymbol{x}_{k+1} \notin W_k = \text{Span}\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k\}$可知 $\boldsymbol{v}_{k+1} \neq \mathbf{0}$. 于是，$\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{k+1}\}$是 $k+1$ 维子空间 W_{k+1}中由非零向量构成的正交集. 因此 $W_{k+1} = \text{Span}\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{k+1}\}$. $k+1 = p$ 时，归纳过程结束. ■

定理 11 表明，$\mathbf{R}^n$的任意非零子空间 W 都有正交基. 这是因为我们总能找到 W 的一组普通基$\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_p\}$(根据 4.5 节定理 11)，并且格拉姆－施密特方法只依赖于有正交基的子空间 W 上正交投影的存在性. 404

6.4.1 标准正交基

一组标准正交基很容易由一组正交基$\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p\}$得到：只需标准化(适当缩放)全体 $\boldsymbol{v}_k$即可. 进行手算时，这种方法比每求出一个 $\boldsymbol{v}_k$就将其标准化更为简单(因为这样可以避免不必要的平方根运算).

【例题 3】 在例题 1 中，我们构造了一组正交基：

$$\boldsymbol{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

则一组标准正交基为：

$$u_1 = \frac{1}{\|v_1\|}v_1 = \frac{1}{\sqrt{45}}\begin{bmatrix}3\\6\\0\end{bmatrix} = \begin{bmatrix}1/\sqrt{5}\\2/\sqrt{5}\\0\end{bmatrix}$$

$$u_2 = \frac{1}{\|v_2\|}v_2 = \begin{bmatrix}0\\0\\1\end{bmatrix}$$

■

6.4.2 矩阵的 QR 分解

若 $m \times n$ 矩阵 A 有线性无关列 $x_1, \cdots, x_n$，那么对它们施以格拉姆－施密特方法（并且进行标准化）相当于按照下述定理对 A 进行分解．这种分解广泛应用于各种计算机算法，例如，求解方程（见 6.5 节）和求本征值（见 5.2 节习题）．

【定理 12】 QR 分解

若 A 是有线性无关列的 $m \times n$ 矩阵，则 A 可以分解为 $A = QR$，其中 Q 是 $m \times n$ 矩阵，且其列构成 Col A 的一组标准正交基，R 是 $n \times n$ 上三角可逆矩阵，且其对角线上元素均为正值．

证明： A 的列构成 Col A 的一组基 $\{x_1, \cdots, x_n\}$．构造 $W = \text{Col } A$ 的一组正交基 $\{u_1, \cdots, u_n\}$，使之满足定理 11 中性质（1）．这组基可以通过运用格拉姆－施密特方法或其他方法得到．令

$$Q = [u_1 \quad u_2 \quad \cdots \quad u_n]$$

405

对 $k = 1, \cdots, n, x_k \in \text{Span}\{x_1, \cdots, x_k\} = \text{Span}\{u_1, \cdots, u_k\}$．于是，存在常数 $r_{1k}, \cdots, r_{kk}$ 使得

$$x_k = r_{1k}u_1 + \cdots + r_{kk}u_k + 0 \cdot u_{k+1} + \cdots + 0 \cdot u_n$$

我们可以假定 $r_{kk} \geqslant 0$（若 $r_{kk} < 0$，将 r_{kk} 和 u_k 同时乘以 -1）．这就表明 x_k 是 Q 中列向量的线性组合，其权重是下列向量的元素：

$$r_k = \begin{bmatrix}r_{1k}\\ \vdots \\ r_{kk} \\ 0 \\ \vdots \\ 0\end{bmatrix}$$

即对 $k = 1, \cdots, n$，有 $x_k = Qr_k$．令 $R = [r_1 \quad r_2 \quad \cdots \quad r_n]$．则

$$A = [x_1 \quad \cdots \quad x_n] = [Qr_1 \quad \cdots \quad Qr_n] = QR$$

A 的各列线性无关这一事实可直接推得 R 可逆（习题 19）．又由于 R 显然是上三角矩阵，其对角线上的元素非负，因而必为正值． ■

【例题 4】 求 $A = \begin{bmatrix}1 & 0 & 0\\1 & 1 & 0\\1 & 1 & 1\\1 & 1 & 1\end{bmatrix}$ 的一个 QR 分解．

解： $\boldsymbol{A}$ 的列就是例题2中的向量 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$ 和 $\boldsymbol{x}_3$．例题2求出了 Col $\boldsymbol{A}$ = Span$\{\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3\}$的一组正交基：

$$\boldsymbol{v}_1=\begin{bmatrix}1\\1\\1\\1\end{bmatrix},\quad \boldsymbol{v}'_2=\begin{bmatrix}-3\\1\\1\\1\end{bmatrix},\quad \boldsymbol{v}_3=\begin{bmatrix}0\\-2/3\\1/3\\1/3\end{bmatrix}$$

令 $\boldsymbol{v}'_3=3\boldsymbol{v}_3$ 以缩放 $\boldsymbol{v}_3$．然后标准化这三个向量，得到 $\boldsymbol{u}_1$，$\boldsymbol{u}_2$ 和 $\boldsymbol{u}_3$，把它们作为 $\boldsymbol{Q}$ 的列：

$$\boldsymbol{Q}=\begin{bmatrix}1/2 & -3/\sqrt{12} & 0\\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6}\\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6}\\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6}\end{bmatrix}$$

由 $\boldsymbol{Q}$ 的构造可知，$\boldsymbol{Q}$ 的前 k 列构成 Span$\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k\}$的一组正交基．根据定理12的证明，存在某个矩阵 $\boldsymbol{R}$ 使得 $\boldsymbol{A}=\boldsymbol{QR}$．为求 $\boldsymbol{R}$，注意到 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q}=\boldsymbol{I}$，这是因为 $\boldsymbol{Q}$ 的各列标准正交．于是

406

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{Q}^{\mathrm{T}}(\boldsymbol{QR})=\boldsymbol{IR}=\boldsymbol{R}$$

且

$$\boldsymbol{R}=\begin{bmatrix}1/2 & 1/2 & 1/2 & 1/2\\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12}\\ 0 & -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6}\end{bmatrix}\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\\1&1&1\end{bmatrix}=\begin{bmatrix}2 & 3/2 & 1\\ 0 & 3/\sqrt{12} & 2/\sqrt{12}\\ 0 & 0 & 2/\sqrt{6}\end{bmatrix}$$ ■

注记

1. 在计算机上运行格拉姆－施密特方法时，随着向量 $\boldsymbol{u}_k$ 被逐个计算出来，舍入误差可能不断累积．对于较大但不相等的 j、k，内积 $\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{u}_k$ 可能不充分接近于零．重新调整计算的次序可以很大程度上弥补这种正交性的损失．[1]然而，我们通常把这种校正了的格拉姆－施密特方法称为基于计算机的另一种QR分解，因为它生成的标准正交基更加精确，尽管这种分解需要的计算量约为其他多数算法的两倍．

2. 为求矩阵 $\boldsymbol{A}$ 的QR分解，计算机算法通常用一系列正交矩阵左乘 $\boldsymbol{A}$，直到将 $\boldsymbol{A}$ 变换为上三角矩阵．这种构造与求 $\boldsymbol{A}$ 的LU分解时用初等矩阵左乘 $\boldsymbol{A}$ 的方法类似．

基础练习

令 $W=\mathrm{Span}\{\boldsymbol{x}_1,\boldsymbol{x}_2\}$，其中 $\boldsymbol{x}_1=\begin{bmatrix}1\\1\\1\end{bmatrix}$，$\boldsymbol{x}_2=\begin{bmatrix}1/3\\1/3\\-2/3\end{bmatrix}$．试构造 W 的一组标准正交基．

---

1. 见 David S. Watkins, *Fundamentals of Matrix Computations* (New York: John Wiley & Sons, 1991), pp. 167－180.

习题 6.4

在习题 1 ~ 6 中，所给集合构成子空间 W 的一组基. 试用格拉姆 - 施密特方法生成 W 的一组正交基.

1. $\begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 8 \\ 5 \\ -6 \end{bmatrix}$

2. $\begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \\ -7 \end{bmatrix}$

3. $\begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix}$

4. $\begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}, \begin{bmatrix} -3 \\ 14 \\ -7 \end{bmatrix}$

5. $\begin{bmatrix} 1 \\ -4 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 7 \\ -7 \\ -4 \\ 1 \end{bmatrix}$

6. $\begin{bmatrix} 3 \\ -1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -5 \\ 9 \\ -9 \\ 3 \end{bmatrix}$

407

7. 求习题 3 中向量所张成子空间的一组标准正交基.
8. 求习题 4 中向量所张成子空间的一组标准正交基.

求习题 9 ~ 12 中各矩阵列空间的一组正交基.

9. $\begin{bmatrix} 3 & -5 & 1 \\ 1 & 1 & 1 \\ -1 & 5 & -2 \\ 3 & -7 & 8 \end{bmatrix}$

10. $\begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & -3 \end{bmatrix}$

11. $\begin{bmatrix} 1 & 2 & 5 \\ -1 & 1 & -4 \\ -1 & 4 & -3 \\ 1 & -4 & 7 \\ 1 & 2 & 1 \end{bmatrix}$

12. $\begin{bmatrix} 1 & 3 & 5 \\ -1 & -3 & 1 \\ 0 & 2 & 3 \\ 1 & 5 & 2 \\ 1 & 5 & 8 \end{bmatrix}$

在习题 13 和 14 中，$\boldsymbol{Q}$ 的列通过对 $\boldsymbol{A}$ 的列施以格拉姆 - 施密特方法而得到. 试求上三角矩阵 $\boldsymbol{R}$ 使 $\boldsymbol{A}=\boldsymbol{QR}$，验证你的答案.

13. $A=\begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix}, \quad Q=\begin{bmatrix} 5/6 & -1/6 \\ 1/6 & 5/6 \\ -3/6 & 1/6 \\ 1/6 & 3/6 \end{bmatrix}$

14. $A=\begin{bmatrix} -2 & 3 \\ 5 & 7 \\ 2 & -2 \\ 4 & 6 \end{bmatrix}, \quad Q=\begin{bmatrix} -2/7 & 5/7 \\ 5/7 & 2/7 \\ 2/7 & -4/7 \\ 4/7 & 2/7 \end{bmatrix}$

15. 求习题 11 中矩阵的 QR 分解.
16. 求习题 12 中矩阵的 QR 分解.

习题 17 和 18 中的所有向量和子空间都在 $\mathbf{R}^n$ 中. 判断各命题的真假，并说明理由.

17. a. 如果 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ 是 W 的一组正交基，则将 $\boldsymbol{v}_3$ 乘上数量 c 可以得到一组新的正交基 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, c\boldsymbol{v}_3\}$.
    b. 格拉姆 - 施密特方法从一个线性无关集 $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_p\}$ 出发，生成一个正交集 $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p\}$，使得对每个 k，向量 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k$ 张成的子空间与 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k$ 张成的子空间相同.
    c. 如果 $\boldsymbol{A}=\boldsymbol{QR}$，其中 $\boldsymbol{Q}$ 有标准正交列，则 $\boldsymbol{R}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}$.
18. a. 如果 $W=\mathrm{Span}\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$，其中 $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$ 线性无关，并且 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ 是 W 中的正交集，则 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ 是 W 的基.
    b. 如果 $\boldsymbol{x}$ 不属于子空间 W，则 $\boldsymbol{x}-\mathrm{proj}_W\,\boldsymbol{x}$ 不是零.
    c. 在 QR 分解 $\boldsymbol{A}=\boldsymbol{QR}$（$\boldsymbol{A}$ 的各列线性无关）当中，$\boldsymbol{Q}$ 的列构成 $\boldsymbol{A}$ 的列空间的一组标准正交基.
19. 设 $\boldsymbol{A}=\boldsymbol{QR}$，其中 $\boldsymbol{Q}$ 和 $\boldsymbol{R}$ 分别是 $m\times n$ 和 $n\times n$ 矩阵. 证明：如果 $\boldsymbol{A}$ 的各列线性无关，则 $\boldsymbol{R}$ 一定可逆. [提示：讨论方程 $\boldsymbol{Rx}=\boldsymbol{0}$ 并利用 $\boldsymbol{A}=\boldsymbol{QR}$ 的事实.]
20. 设 $\boldsymbol{A}=\boldsymbol{QR}$，其中 $\boldsymbol{R}$ 是可逆矩阵. 证明：如果 $\boldsymbol{A}$ 和 $\boldsymbol{Q}$ 有相同列空间. [提示：给定 $\boldsymbol{y}\in\mathrm{Col}\,\boldsymbol{A}$，证明存在 $\boldsymbol{x}$ 使 $\boldsymbol{y}=\boldsymbol{Qx}$. 另外，给定 $\boldsymbol{y}\in\mathrm{Col}\,\boldsymbol{Q}$，证明存在 $\boldsymbol{x}$ 使 $\boldsymbol{y}=\boldsymbol{Ax}$.]

21. 设 $A=QR$ 同定理12，说明如何求正交 $m\times m$(方)阵 Q_1 以及可逆 $n\times n$ 上三角矩阵 R 使得

$$A=Q_1\begin{bmatrix}R\\0\end{bmatrix}$$

当 rank $A=n$ 时，MATLAB 的 qr 命令给出这一“完全”QR 分解.

22. 令 $u_1,\cdots,u_p$ 是 $\mathbb{R}^n$ 子空间 W 的一组正交基，T：$\mathbb{R}^n\to\mathbb{R}^n$ 定义为 $T(x)=\text{proj}_W x$. 证明 T 是线性变换.

23. 设 $A=QR$ 是 $m\times n$ 矩阵 A（其中各列线性无关)的 QR 分解. 将 A 划分成 $[A_1\quad A_2]$，其中 A_1 含 p 列. 说明如何得到 A_1 的 QR 分解，并解释所求的分解为什么具有相应的性质.

24. [M]利用格拉姆-施密特方法，仿照例题2生成下述矩阵列空间的一组正交基：

$$A=\begin{bmatrix}-10&13&7&-11\\2&1&-5&3\\-6&3&13&-3\\16&-16&-2&5\\2&1&-5&-7\end{bmatrix}$$

25. [M]利用本节方法，给出习题24中矩阵的 QR 分解.

26. [M]格拉姆-施密特方法在处理矩阵时比求正交向量更为有效. 从定理11中的 $x_1,\cdots,x_p$ 出发，令 $A=[x_1\cdots x_p]$. 设 Q 是 $n\times k$ 矩阵且其列构成 A 的前 k 列所张成子空间 W_k 的一组标准正交基. 则对 $x\in\mathbb{R}^n$，$QQ^{\mathrm{T}}x$ 是 x 在 W_k 上的正交投影(6.3节定理10). 若 x_{k+1} 是 A 的第 $k+1$ 列，则定理11证明中的(2)式变为：

$$v_{k+1}=x_{k+1}-Q(Q^{\mathrm{T}}x_{k+1})$$

(上面的圆括号可以减少运算的次数.)

令 $u_{k+1}=v_{k+1}/\|v_{k+1}\|$. 下一步 Q 更新为 $[Q\quad u_{k+1}]$. 利用这一算法计算习题24中矩阵的 QR 分解. 写出你所用的按键和命令.

408

基础练习答案

令 $v_1=x_1=\begin{bmatrix}1\\1\\1\end{bmatrix}$，且 $v_2=x_2-\dfrac{x_2\cdot v_1}{v_1\cdot v_1}v_1=x_2-0v_1=x_2$. 于是 $\{x_1,x_2\}$ 正交. 下面只需标准化向量即可. 令

$$u_1=\frac{1}{\|v_1\|}v_1=\frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$$

我们不直接对向量 v_2 进行标准化，而代之以 $v'_2=3v_2$：

$$u_2=\frac{1}{\|v'_2\|}v'_2=\frac{1}{\sqrt{1^2+1^2+(-2)^2}}\begin{bmatrix}1\\1\\-2\end{bmatrix}=\begin{bmatrix}1/\sqrt{6}\\1/\sqrt{6}\\-2/\sqrt{6}\end{bmatrix}$$

这样，$\{u_1,u_2\}$ 构成 W 的一组标准正交基.

6.5　最小二乘问题

本章实例介绍了一个无解的大型问题 $Ax=b$. 实际应用中，虽然系数矩阵不一定

很庞大，但还是常出现不相容的系统．当系统要求一个解但却无解时，最好的方法是求 $\boldsymbol{x}$ 使 $A\boldsymbol{x}$ 尽可能地接近 $\boldsymbol{b}$.

考虑 $A\boldsymbol{x}$ 作为 $\boldsymbol{b}$ 的逼近．$\boldsymbol{b}$ 到 $A\boldsymbol{x}$ 的距离 $\|\boldsymbol{b}-A\boldsymbol{x}\|$ 越小，这个逼近就越好．**一般最小二乘问题**(general least-squares problem)就是求 $\boldsymbol{x}$ 使 $\|\boldsymbol{b}-A\boldsymbol{x}\|$ 尽可能最小．“最小二乘”一词来源于 $\|\boldsymbol{b}-A\boldsymbol{x}\|$ 是一个平方和的平方根这一事实.

【定义】 若 A 是 $m\times n$ 矩阵且向量 $\boldsymbol{b}\in\mathbb{R}^m$，$A\boldsymbol{x}=\boldsymbol{b}$ 的一个**最小二乘解**(least-squares solution)是 $\mathbb{R}^n$ 中的向量 $\hat{\boldsymbol{x}}$，使得对全体 $\boldsymbol{x}\in\mathbb{R}^n$，

$$\|\boldsymbol{b}-A\hat{\boldsymbol{x}}\|\leqslant\|\boldsymbol{b}-A\boldsymbol{x}\|$$

409

最小二乘问题的一个重要特点是，不论我们选取什么 $\boldsymbol{x}$，向量 $A\boldsymbol{x}$ 必定属于 A 的列空间 Col A．因此我们希望找到 $\boldsymbol{x}$，使得 $A\boldsymbol{x}$ 是 Col A 中与 $\boldsymbol{b}$ 距离最近的点，如图 6-24所示．(当然，如果 $\boldsymbol{b}$ 恰好属于 Col A，那么存在某个 $\boldsymbol{x}$ 使得 $\boldsymbol{b}=A\boldsymbol{x}$，这个 $\boldsymbol{x}$ 就是一个“最小二乘解”.)

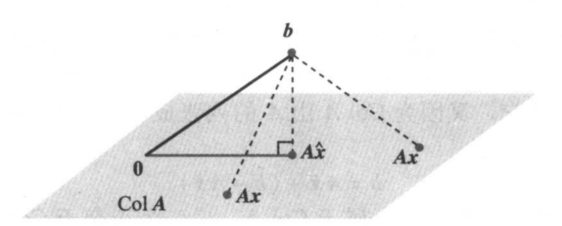


图 6-24 向量 $\boldsymbol{b}$ 到 $A\hat{\boldsymbol{x}}$ 的距离比到其他 $A\boldsymbol{x}$ 的距离都近

6.5.1 一般最小二乘问题的解

给定 A，$\boldsymbol{b}$ 同上，对子空间 Col A 运用 6.3 节的最佳逼近定理．令

$$\hat{\boldsymbol{b}}=\text{proj}_{\text{Col}\,A}\,\boldsymbol{b}$$

因为 $\hat{\boldsymbol{b}}$ 属于 A 的列空间，方程 $A\boldsymbol{x}=\hat{\boldsymbol{b}}$ 是相容的，并且存在 $\hat{\boldsymbol{x}}\in\mathbb{R}^n$ 使得

$$A\hat{\boldsymbol{x}}=\hat{\boldsymbol{b}} \tag{1}$$

由于 $\hat{\boldsymbol{b}}$ 是 Col A 中与 $\boldsymbol{b}$ 距离最近的点，向量 $\hat{\boldsymbol{x}}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解当且仅当 $\hat{\boldsymbol{x}}$ 满足(1)式．$\mathbb{R}^n$ 中这样的 $\hat{\boldsymbol{x}}$ 是权重的列表，它们决定了由 A 中的列构造 $\boldsymbol{b}$ 的方式．如图 6-25．[如果(1)式含有自由变量，则(1)有多个解.]

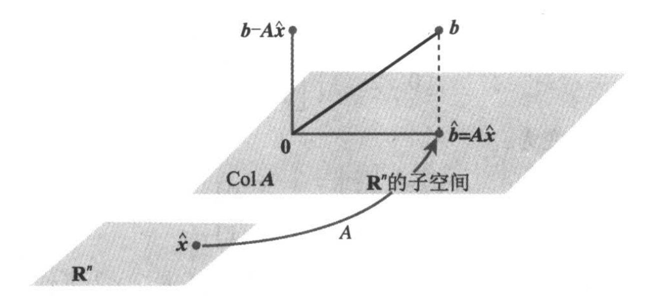


图 6-25 最小二乘解 $\hat{\boldsymbol{x}}$ 属于 $\mathbb{R}^n$

假设 $\hat{\boldsymbol{x}}$ 满足 $\boldsymbol{A}\hat{\boldsymbol{x}}=\hat{\boldsymbol{b}}$. 根据6.3节正交分解定理，投影 $\hat{\boldsymbol{b}}$ 具有性质：$\boldsymbol{b}-\hat{\boldsymbol{b}}$ 正交于
410
Col $\boldsymbol{A}$，于是 $\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{x}}$ 与 $\boldsymbol{A}$ 的每一列都正交. 若 $\boldsymbol{a}_j$ 是 $\boldsymbol{A}$ 的某一列，则 $\boldsymbol{a}_j\cdot(\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{x}})=0$ 且 $\boldsymbol{a}_j^{\mathrm{T}}(\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{x}})=0$. 由于每个 $\boldsymbol{a}_j^{\mathrm{T}}$ 都是 $\boldsymbol{A}^{\mathrm{T}}$ 的一行，

$$\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{x}})=\boldsymbol{0} \tag{2}$$

(这个式子也可由6.1节定理3推出). 这样，

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\hat{\boldsymbol{x}}=\boldsymbol{0}$$

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\hat{\boldsymbol{x}}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$$

这一计算过程表明 $\boldsymbol{Ax}=\boldsymbol{b}$ 的每个最小二乘解都满足方程：

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} \tag{3}$$

矩阵方程(3)表示的一组方程称作 $\boldsymbol{Ax}=\boldsymbol{b}$ 的**正规方程**(normal equation). (3)的解常记作 $\hat{\boldsymbol{x}}$.

【定理13】 $\boldsymbol{Ax}=\boldsymbol{b}$ 最小二乘解的集合等于正规方程 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$ 的解集，并且非空.

证明：如上文所述，最小二乘解的集合非空，且每个最小二乘解 $\hat{\boldsymbol{x}}$ 都满足正规方程. 反过来，设 $\hat{\boldsymbol{x}}$ 满足 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$. 则 $\hat{\boldsymbol{x}}$ 满足上文(2)式，这就表明 $\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{x}}$ 与 $\boldsymbol{A}^{\mathrm{T}}$ 的行正交，于是与 $\boldsymbol{A}$ 的列正交. 又因为 Col $\boldsymbol{A}$ 由 $\boldsymbol{A}$ 的列张成，向量 $\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{x}}$ 与整个 Col $\boldsymbol{A}$ 正交. 所以方程

$$\boldsymbol{b}=\boldsymbol{A}\hat{\boldsymbol{x}}+(\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{x}})$$

将 $\boldsymbol{b}$ 分解成两向量之和，其中一个属于 Col $\boldsymbol{A}$，另一个正交于 Col $\boldsymbol{A}$. 根据正交分解的唯一性，$\boldsymbol{A}\hat{\boldsymbol{x}}$ 一定是 $\boldsymbol{b}$ 在 Col $\boldsymbol{A}$ 上的正交投影，即 $\boldsymbol{A}\hat{\boldsymbol{x}}=\hat{\boldsymbol{b}}$，且 $\hat{\boldsymbol{x}}$ 是一个最小二乘解. ■

【例题1】　求不相容系统 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小二乘解，其中

$$\boldsymbol{A}=\begin{bmatrix}4&0\\0&2\\1&1\end{bmatrix},\quad \boldsymbol{b}=\begin{bmatrix}2\\0\\11\end{bmatrix}$$

解：为利用(3)，首先计算：

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{bmatrix}4&0&1\\0&2&1\end{bmatrix}\begin{bmatrix}4&0\\0&2\\1&1\end{bmatrix}=\begin{bmatrix}17&1\\1&5\end{bmatrix}$$

411

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}=\begin{bmatrix}4&0&1\\0&2&1\end{bmatrix}\begin{bmatrix}2\\0\\11\end{bmatrix}=\begin{bmatrix}19\\11\end{bmatrix}$$

于是方程 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$ 变为：

$$\begin{bmatrix}17&1\\1&5\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}19\\11\end{bmatrix}$$

我们可以利用行变换来求解这个方程组，不过由于 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是 2×2 可逆矩阵，直接计算下式会更快些：

$$(A^{\mathrm{T}}A)^{-1}=\frac{1}{84}\begin{bmatrix}5 & -1\\ -1 & 17\end{bmatrix}$$

接下来可按下列步骤求解 $A^{\mathrm{T}}Ax=A^{\mathrm{T}}b$：

$$\hat{x}=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b=\frac{1}{84}\begin{bmatrix}5 & -1\\ -1 & 17\end{bmatrix}\begin{bmatrix}19\\ 11\end{bmatrix}=\frac{1}{84}\begin{bmatrix}84\\ 168\end{bmatrix}=\begin{bmatrix}1\\ 2\end{bmatrix}$$ ■

在许多实际计算中，$A^{\mathrm{T}}A$ 都可逆，但也不全是如此. 下面例题所涉及的矩阵源于统计学中所谓的方差分析问题，它就是一个例外.

【例题 2】 求 $Ax=b$ 的最小二乘解，其中

$$A=\begin{bmatrix}1&1&0&0\\1&1&0&0\\1&0&1&0\\1&0&1&0\\1&0&0&1\\1&0&0&1\end{bmatrix},\quad b=\begin{bmatrix}-3\\-1\\0\\2\\5\\1\end{bmatrix}$$

解：计算

$$A^{\mathrm{T}}A=\begin{bmatrix}1&1&1&1&1&1\\1&1&0&0&0&0\\0&0&1&1&0&0\\0&0&0&0&1&1\end{bmatrix}\begin{bmatrix}1&1&0&0\\1&1&0&0\\1&0&1&0\\1&0&1&0\\1&0&0&1\\1&0&0&1\end{bmatrix}=\begin{bmatrix}6&2&2&2\\2&2&0&0\\2&0&2&0\\2&0&0&2\end{bmatrix}$$

$$A^{\mathrm{T}}b=\begin{bmatrix}1&1&1&1&1&1\\1&1&0&0&0&0\\0&0&1&1&0&0\\0&0&0&0&1&1\end{bmatrix}\begin{bmatrix}-3\\-1\\0\\2\\5\\1\end{bmatrix}=\begin{bmatrix}4\\-4\\2\\6\end{bmatrix}$$

$A^{\mathrm{T}}Ax=A^{\mathrm{T}}b$ 的增广矩阵是

$$\begin{bmatrix}6&2&2&2&4\\2&2&0&0&-4\\2&0&2&0&2\\2&0&0&2&6\end{bmatrix}\sim\begin{bmatrix}1&0&0&1&3\\0&1&0&-1&-5\\0&0&1&-1&-2\\0&0&0&0&0\end{bmatrix}$$

412

通解为 $x_1=3-x_4$，$x_2=-5+x_4$，$x_3=-2+x_4$，x_4是自由变量. 所以 $Ax=b$ 的最小二乘通解具有如下形式：

$$\hat{x}=\begin{bmatrix}3\\-5\\-2\\0\end{bmatrix}+x_4\begin{bmatrix}-1\\1\\1\\1\end{bmatrix}$$ ■

下面的定理就 $\boldsymbol{Ax}=\boldsymbol{b}$ 何时有唯一解给出了一个实用的判别方法(当然，正交投影总是唯一的).

【定理 14】 矩阵 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 可逆当且仅当 $\boldsymbol{A}$ 的各列线性无关. 此时，方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 有唯一的最小二乘解 $\hat{\boldsymbol{x}}$，且 $\hat{\boldsymbol{x}}$ 由下式给出：

$$\hat{\boldsymbol{x}}=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} \tag{4}$$

习题 19 ~ 21 给出了定理 14 证明的轮廓，定理证明的同时也复习了第 4 章的概念. $\hat{\boldsymbol{x}}$ 的计算公式(4)主要具有较强的理论意义，当 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是 2×2 可逆矩阵时，也可以用于手算.

使用最小二乘解 $\hat{\boldsymbol{x}}$ 来构造 $\boldsymbol{b}$ 的一个逼近 $\boldsymbol{A}\hat{\boldsymbol{x}}$ 时，$\boldsymbol{b}$ 到 $\boldsymbol{A}\hat{\boldsymbol{x}}$ 的距离被称作这个逼近的**最小二乘误差**（least-squares errror).

【例题 3】 给定 $\boldsymbol{A}$，$\boldsymbol{b}$ 如例题 1，确定 $\boldsymbol{Ax}=\boldsymbol{b}$ 最小二乘解的最小二乘误差.

解： 由例题可知：

$$\boldsymbol{b}=\begin{bmatrix}2\\0\\11\end{bmatrix}\text{和 }\boldsymbol{A}\hat{\boldsymbol{x}}=\begin{bmatrix}4&0\\0&2\\1&1\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix}=\begin{bmatrix}4\\4\\3\end{bmatrix}$$

于是

$$\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{x}}=\begin{bmatrix}2\\0\\11\end{bmatrix}-\begin{bmatrix}4\\4\\3\end{bmatrix}=\begin{bmatrix}-2\\-4\\8\end{bmatrix}$$

且

$$\|\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{x}}\|=\sqrt{(-2)^2+(-4)^2+8^2}=\sqrt{84}$$

最小二乘误差是 $\sqrt{84}$. 对任意 $\boldsymbol{x}\in\mathbf{R}^n$，$\boldsymbol{b}$ 到向量 $\boldsymbol{A}\hat{\boldsymbol{x}}$ 的距离至少是 $\sqrt{84}$，如图 6-26 所示. 注意：最小二乘解 $\hat{\boldsymbol{x}}$ 本身并未出现在图中.

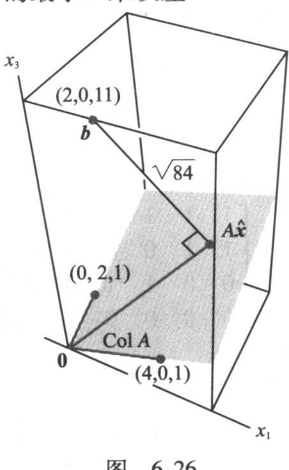


图　6-26

413

■

6.5.2　最小二乘解的另一种算法

下面的例题演示了当 $\boldsymbol{A}$ 中各列正交时，如何求解 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小二乘解. 这类矩阵经常出现在下节要讨论的线性回归问题当中.

【例题 4】 求 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小二乘解，其中

$$\boldsymbol{A}=\begin{bmatrix}1&-6\\1&-2\\1&1\\1&7\end{bmatrix},\quad \boldsymbol{b}=\begin{bmatrix}-1\\2\\1\\6\end{bmatrix}$$

解： 因为 $\boldsymbol{A}$ 的列 $\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2$ 正交，$\boldsymbol{b}$ 在 Col $\boldsymbol{A}$ 上的正交投影由下式给出：

$$\hat{\boldsymbol{b}}=\frac{\boldsymbol{b}\cdot\boldsymbol{a}_1}{\boldsymbol{a}_1\cdot\boldsymbol{a}_1}\boldsymbol{a}_1+\frac{\boldsymbol{b}\cdot\boldsymbol{a}_2}{\boldsymbol{a}_2\cdot\boldsymbol{a}_2}\boldsymbol{a}_2=\frac{8}{4}\boldsymbol{a}_1+\frac{45}{90}\boldsymbol{a}_2=\begin{bmatrix}2\\2\\2\\2\end{bmatrix}+\begin{bmatrix}-3\\-1\\1/2\\7/2\end{bmatrix}=\begin{bmatrix}-1\\1\\5/2\\11/2\end{bmatrix} \tag{5}$$

既然已知 $\hat{b}$，我们便可求解 $A\hat{x}=\hat{b}$. 这很容易，因为我们知道，为得到 $\hat{b}$，需在 A 的各列上添加何种权重. 由(5)显然有：

$$\hat{x}=\begin{bmatrix}8/4\\45/90\end{bmatrix}=\begin{bmatrix}2\\1/2\end{bmatrix}$$ ■

在某些情况下，最小二乘问题的正规方程可能是病态的；即，A^TA 中的元素在计算过程中很小误差有可能导致解 $\hat{x}$ 的较大误差. 如果 A 中各列线性无关，最小二乘解往往可以通过 A 的 QR 分解(6.4 中有介绍)更可靠地计算出来.[1]

【定理 15】 给定各列线性无关的 $m\times n$ 矩阵 A，设 $A=QR$ 同定理 12 中的 QR 分解. 则对每个 $b\in\mathbb{R}^m$，方程 $Ax=b$ 有唯一的最小二乘解，且由下式给出：

$$\hat{x}=R^{-1}Q^Tb \tag{6}$$

414

证明： 令 $\hat{x}=R^{-1}Q^Tb$. 则

$$A\hat{x}=QR\hat{x}=QRR^{-1}Q^Tb=QQ^Tb$$

根据定理 12，Q 的列构成 Col A 的一组正交基. 于是由定理 10 可知，QQ^Tb 是 b 在 Col A 上的正交投影 $\hat{b}$. 则 $A\hat{x}=\hat{b}$，这就表明 $\hat{x}$ 是 $Ax=b$ 的最小二乘解. $\hat{x}$ 的唯一性可由定理 14 推得. ■

注记

由于定理 15 中的 R 是上三角矩阵，$\hat{x}$ 可以作为下面方程的解来计算：

$$Rx=Q^Tb \tag{7}$$

利用回代或行变换方法求解(7)比计算 R^{-1} 和利用(6)要快得多.

【例题 5】 求 $Ax=b$ 的最小二乘解，其中

$$A=\begin{bmatrix}1&3&5\\1&1&0\\1&1&2\\1&3&3\end{bmatrix},\quad b=\begin{bmatrix}3\\5\\7\\-3\end{bmatrix}$$

解： A 的 QR 分解可由 6.4 节得到：

$$A=QR=\begin{bmatrix}1/2&1/2&1/2\\1/2&-1/2&-1/2\\1/2&-1/2&1/2\\1/2&1/2&-1/2\end{bmatrix}\begin{bmatrix}2&4&5\\0&2&3\\0&0&2\end{bmatrix}$$

于是

$$Q^Tb=\begin{bmatrix}1/2&1/2&1/2&1/2\\1/2&-1/2&-1/2&1/2\\1/2&-1/2&1/2&-1/2\end{bmatrix}\begin{bmatrix}3\\5\\7\\-3\end{bmatrix}=\begin{bmatrix}6\\-6\\4\end{bmatrix}$$

1. QR 方法与标准正规方程方法的比较见 G. Golub, C. Van Loan, *Matrix Computations*, 3rd ed. (Baltimore: Johns Hopkins Press. 1996), pp. 230－231.

最小二乘解 $\hat{x}$ 满足 $Rx=Q^{T}b$，即

$$\begin{bmatrix}2&4&5\\0&2&3\\0&0&2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}6\\-6\\4\end{bmatrix}$$

415 容易求解这个方程，得到 $\hat{x}=\begin{bmatrix}10\\-6\\2\end{bmatrix}$. ■

基础练习

1. 令 $A=\begin{bmatrix}1&-3&-3\\1&5&1\\1&7&2\end{bmatrix}$, $b=\begin{bmatrix}5\\-3\\-5\end{bmatrix}$. 求 $Ax=b$ 的最小二乘解，并计算相应的最小二乘误差.
2. 当 b 与 A 的各列正交时，关于 $Ax=b$ 的最小二乘解，你能做出什么结论?

习题　6.5

在习题 1 ~ 4 中，按下列步骤求 $Ax=b$ 的最小二乘解：(a)构造 $\hat{x}$ 的正规方程；(b)求解 $\hat{x}$.

1. $A=\begin{bmatrix}-1&2\\2&-3\\-1&3\end{bmatrix}$, $b=\begin{bmatrix}4\\1\\2\end{bmatrix}$

2. $A=\begin{bmatrix}2&1\\-2&0\\2&3\end{bmatrix}$, $b=\begin{bmatrix}-5\\8\\1\end{bmatrix}$

3. $A=\begin{bmatrix}1&-2\\-1&2\\0&3\\2&5\end{bmatrix}$, $b=\begin{bmatrix}3\\1\\-4\\2\end{bmatrix}$

4. $A=\begin{bmatrix}1&3\\1&-1\\1&1\end{bmatrix}$, $b=\begin{bmatrix}5\\1\\0\end{bmatrix}$

在习题 5 和 6 中，求 $Ax=b$ 的全体最小二乘解.

5. $A=\begin{bmatrix}1&1&0\\1&1&0\\1&0&1\\1&0&1\end{bmatrix}$, $b=\begin{bmatrix}1\\3\\8\\2\end{bmatrix}$

6. $A=\begin{bmatrix}1&1&0\\1&1&0\\1&1&0\\1&0&1\\1&0&1\\1&0&1\end{bmatrix}$, $b=\begin{bmatrix}7\\2\\3\\6\\5\\4\end{bmatrix}$

7. 计算习题 3 中最小二乘解的最小二乘误差.
8. 计算习题 4 中最小二乘解的最小二乘误差.

在习题 9 ~ 12 中，求(a)b 在 Col A 上的正交投影；(b)$Ax=b$ 的最小二乘解.

9. $A=\begin{bmatrix}1&5\\3&1\\-2&4\end{bmatrix}$, $b=\begin{bmatrix}4\\-2\\-3\end{bmatrix}$

10. $A=\begin{bmatrix}1&2\\-1&4\\1&2\end{bmatrix}$, $b=\begin{bmatrix}3\\-1\\5\end{bmatrix}$

11. $A=\begin{bmatrix}4&0&1\\1&-5&1\\6&1&0\\1&-1&-5\end{bmatrix}$, $b=\begin{bmatrix}9\\0\\0\\0\end{bmatrix}$

12. $A=\begin{bmatrix}1&1&0\\1&0&-1\\0&1&1\\-1&1&-1\end{bmatrix}$, $b=\begin{bmatrix}2\\5\\6\\6\end{bmatrix}$

13. 令 $A=\begin{bmatrix}3&4\\-2&1\\3&4\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}11\\-9\\5\end{bmatrix}$, $\boldsymbol{u}=\begin{bmatrix}5\\-1\end{bmatrix}$以及 $\boldsymbol{v}=\begin{bmatrix}5\\-2\end{bmatrix}$. 计算 $A\boldsymbol{u}$ 和 $A\boldsymbol{v}$，并与 $\boldsymbol{b}$ 作比较. $\boldsymbol{u}$ 是否可能是 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解？（不计算最小二乘解，回答这一问题.）

14. 令 $A=\begin{bmatrix}2&1\\-3&-4\\3&2\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}5\\4\\4\end{bmatrix}$, $\boldsymbol{u}=\begin{bmatrix}4\\-5\end{bmatrix}$以及 $\boldsymbol{v}=\begin{bmatrix}6\\-5\end{bmatrix}$. 计算 $A\boldsymbol{u}$ 和 $A\boldsymbol{v}$，并与 $\boldsymbol{b}$ 作比较. 是否可能 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 当中至少有一个是 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解？（不计算最小二乘解，回答这一问题.）

416

在习题 15 和 16 中，利用分解 $A=QR$，求 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解.

15. $A=\begin{bmatrix}2&3\\2&4\\1&1\end{bmatrix}=\begin{bmatrix}2/3&-1/3\\2/3&2/3\\1/3&-2/3\end{bmatrix}\begin{bmatrix}3&5\\0&1\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}7\\3\\1\end{bmatrix}$

16. $A=\begin{bmatrix}1&-1\\1&4\\1&-1\\1&4\end{bmatrix}=\begin{bmatrix}1/2&-1/2\\1/2&1/2\\1/2&-1/2\\1/2&1/2\end{bmatrix}\begin{bmatrix}2&3\\0&5\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}-1\\6\\5\\7\end{bmatrix}$

在习题 17 和 18 中，A 是 $m\times n$ 矩阵且 $\boldsymbol{b}\in\mathbf{R}^n$. 判断各命题的真假，并说明理由.

17. a. 一般的最小二乘问题是求 $\boldsymbol{x}$，使得 $A\boldsymbol{x}$ 尽可能接近 $\boldsymbol{b}$.
    b. $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解是满足 $A\hat{\boldsymbol{x}}=\hat{\boldsymbol{b}}$ 的向量 $\hat{\boldsymbol{x}}$，其中 $\hat{\boldsymbol{b}}$ 是 $\boldsymbol{b}$ 在 Col A 上的正交投影.
    c. $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解是使得对任意 $\boldsymbol{x}\in\mathbf{R}^n$ 都有 $\|\boldsymbol{b}-A\boldsymbol{x}\|\leq\|\boldsymbol{b}-A\hat{\boldsymbol{x}}\|$ 的向量 $\hat{\boldsymbol{x}}$.
    d. $A^{\mathrm{T}}A\boldsymbol{x}=A^{\mathrm{T}}\boldsymbol{b}$ 的任意解都是 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解.
    e. 若 A 的各列线性无关，则方程 $A\boldsymbol{x}=\boldsymbol{b}$ 恰有一个最小二乘解.
18. a. 如果 $\boldsymbol{b}$ 属于 A 的列空间，则 $A\boldsymbol{x}=\boldsymbol{b}$ 的每个解都是最小二乘解.
    b. $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解是 A 的列空间中与 $\boldsymbol{b}$ 距离最近的点.
    c. $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解是若干权重的列表，这些权重施加在 A 的列上就得到 $\boldsymbol{b}$ 在 Col A 上的正交投影.
    d. 若 $\hat{\boldsymbol{x}}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解，则 $\hat{\boldsymbol{x}}=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b}$.
    e. 用正规方程来计算最小二乘解总是可靠的.
    f. 如果 A 有 QR 分解，比如说 $A=QR$，那么求 $A\boldsymbol{x}=\boldsymbol{b}$ 最小二乘解的最好方法是计算 $\hat{\boldsymbol{x}}=R^{-1}Q^{\mathrm{T}}\boldsymbol{b}$.
19. 设 A 是 $m\times n$ 矩阵. 利用下列步骤证明：$\mathbf{R}^n$ 中的向量 $\boldsymbol{x}$ 满足 $A\boldsymbol{x}=\boldsymbol{0}$ 当且仅当 $A^{\mathrm{T}}A\boldsymbol{x}=\boldsymbol{0}$. 这就说明 Nul A = Nul $A^{\mathrm{T}}A$.
    a. 证明：若 $A\boldsymbol{x}=\boldsymbol{0}$，则 $A^{\mathrm{T}}A\boldsymbol{x}=\boldsymbol{0}$.
    b. 设 $A^{\mathrm{T}}A\boldsymbol{x}=\boldsymbol{0}$. 说明为什么有 $\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{x}=0$，并以此证明 $A\boldsymbol{x}=\boldsymbol{0}$.
20. 设 A 是使 $A^{\mathrm{T}}A$ 可逆的 $m\times n$ 矩阵. 证明 A 中各列线性无关. [注意：不能假定 A 是可逆的，它也不一定是方阵.]
21. 设 A 是各列线性无关的 $m\times n$ 矩阵. [注意：A 不需要是方阵.]
    a. 利用习题 19 证明 $A^{\mathrm{T}}A$ 是可逆矩阵.
    b. 说明为什么 A 的行数一定大于或等于列数.
    c. 确定 A 的秩.
22. 利用习题 19 证明：rank $A^{\mathrm{T}}A$ = rank A. [提示：$A^{\mathrm{T}}A$ 有多少列？这与 $A^{\mathrm{T}}A$ 的秩有何联系?]

23. 设 A 是各列线性无关的 $m \times n$ 矩阵且 $b \in \mathbf{R}^n$. 利用正规方程给出 $\hat{b}$ 的计算公式，这里 $\hat{b}$ 是 b 在 Col A 上的投影[提示：先求 $\hat{x}$. 所求的公式中不需要用 Col A 的正交基].
24. 当 A 的各列标准正交时，找出 $Ax = b$ 最小二乘解的公式.
25. 求下列方程组的全部最小二乘解：

$$x + y = 2$$
$$x + y = 4$$

26. [M]4.8 节例题 3 演示了一种低通线性滤波器，它将信号 $\{y_k\}$ 转换为 $\{y_{k+1}\}$，并且将高频信号 $\{w_k\}$ 转化为零信号，其中 $y_k = \cos(\pi k/4)$ 且 $w_k = \cos(3\pi k/4)$. 下面的算法将设计一种性能类似的滤波器，其滤波方程为：

$$\text{对任意 } k,\ a_0 y_{k+2} + a_1 y_{k+1} + a_2 y_k = z_k \tag{8}$$

由于信号是周期性的，且周期为 8，因此只需要讨论方程(8)在 $k = 0, \cdots, 7$ 时的情形. 前面所描述的对两种信号的操作可以写成下列两个含 8 个方程的方程组：

$$\begin{array}{c} \\ k=0 \\ k=1 \\ \vdots \\ \\ \\ \\ \\ k=7 \end{array}
\begin{array}{c} \begin{array}{ccc} y_{k+2} & y_{k+1} & y_k \end{array} \\ \begin{bmatrix} 0 & 0.7 & 1 \\ -0.7 & 0 & 0.7 \\ -1 & -0.7 & 0 \\ -0.7 & -1 & -0.7 \\ 0 & -0.7 & -1 \\ 0.7 & 0 & -0.7 \\ 1 & 0.7 & 0 \\ 0.7 & 1 & 0.7 \end{bmatrix} \end{array}
\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} =
\begin{array}{c} y_{k+1} \\ \begin{bmatrix} 0.7 \\ 0 \\ -0.7 \\ -1 \\ -0.7 \\ 0 \\ 0.7 \\ 1 \end{bmatrix} \end{array};$$

$$\begin{array}{c} \\ k=0 \\ k=1 \\ \vdots \\ \\ \\ \\ \\ k=7 \end{array}
\begin{array}{c} \begin{array}{ccc} w_{k+2} & w_{k+1} & w_k \end{array} \\ \begin{bmatrix} 0 & -0.7 & 1 \\ 0.7 & 0 & -0.7 \\ -1 & 0.7 & 0 \\ 0.7 & -1 & 0.7 \\ 0 & 0.7 & -1 \\ -0.7 & 0 & 0.7 \\ 1 & -0.7 & 0 \\ -0.7 & 1 & -0.7 \end{bmatrix} \end{array}
\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} =
\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

将它们写成一个方程 $Ax = b$，其中 A 是上面两个系数矩阵构成的 16×3 矩阵，b 是 $\mathbf{R}^n$ 中的向量，由上面两方程的右端部分构成. 试求由 $Ax = b$ 的最小二乘解所确定的 a_0，a_1 和 a_2.（前文数据中的 0.7 是 $\sqrt{2}/2$ 的一个逼近值，用以说明实际问题中是如何展开计算的. 如果代之以 0.707，计算所得的滤波系数与精确的计算结果 $\sqrt{2}/4$，$1/2$ 和 $\sqrt{2}/4$ 相比，至少有七位相同的小数位.）

基础练习答案

1. 首先，计算

$$A^{\mathrm{T}}A = \begin{bmatrix} 1 & 1 & 1 \\ -3 & 5 & 7 \\ -3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 9 & 0 \\ 9 & 83 & 28 \\ 0 & 28 & 14 \end{bmatrix}$$

$$A^{\mathrm{T}}b = \begin{bmatrix} 1 & 1 & 1 \\ -3 & 5 & 7 \\ -3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix} = \begin{bmatrix} -3 \\ -65 \\ -28 \end{bmatrix}$$

接下来，行化简正规方程 $A^TAx = A^Tb$ 的增广矩阵：

$$\begin{bmatrix} 3 & 9 & 0 & -3 \\ 9 & 83 & 28 & -65 \\ 0 & 28 & 14 & -28 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 0 & -1 \\ 0 & 56 & 28 & -56 \\ 0 & 28 & 14 & -28 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & -3/2 & 2 \\ 0 & 1 & 1/2 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

最小二乘通解是 $x_1 = 2 + \frac{3}{2}x_3$，$x_2 = -1 - \frac{1}{2}x_3$，其中 x_3 是自由变量. 对于某个特解，例如取 $x_3 = 0$，可得

$$\hat{x} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$$

为求出最小二乘误差，计算

$$\hat{b} = A\hat{x} = \begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix}$$

最后有 $\hat{b} = b$，于是 $\|b - \hat{b}\| = 0$. 因为 b 恰好属于 Col A，所以最小二乘误差是零.

2. 如果 b 与 A 的各列正交，则 b 在 A 的列空间上的投影是 $\mathbf{0}$. 此时，$Ax = b$ 的最小二乘解 $\hat{x}$ 满足 $A\hat{x} = \mathbf{0}$.

418

6.6 线性模型中的应用

科学和工程中的一项普遍任务是分析和理解多个变量之间的关系. 本节讨论了一系列例子，其中的数据被用于构造或验证一类公式，这些公式将变量表示成其他变量的函数. 每个例子里的问题最后都相当于求解一个最小二乘问题.

为方便将这些讨论应用到今后遇到的实际问题当中，我们采用科学和工程数据统计分析中的常用记号：将 $Ax = b$ 记为 $X\beta = y$，其中 X，β 和 y 分别称作**设计矩阵**(design matrix)、**参数向量**(parameter vector)和**观测向量**(observation vector).

6.6.1 最小二乘直线

两个变量 x 和 y 之间最简单的关系是线性关系 $y = \beta_0 + \beta_1 x$.[1]实验数据所产生的点 $(x_1, y_1), \cdots, (x_n, y_n)$ 绘成图像时常近似地分布在某条直线附近. 我们希望确定参数 β_0 和 β_1，以使这条直线尽可能“接近”这些点.

假设 β_0 和 β_1 都固定，考虑图6-27 中的直线 $y = \beta_0 + \beta_1 x$. 与每个数据点 (x_j, y_j) 相对应，直线上都有一个具有相同 x-坐标的点 $(x_j, \beta_0 + \beta_1 x_j)$. 我们称 y_j 为 y 的观测值，称 $\beta_0 + \beta_1 x_j$ 为 y 的预测值(由直线确定). y 的一个观测值与对应预测值的差称为一个剩余.

有多种方式用于度量直线有多“接近”这些数据. 常用的一种(主要因为其数学算法简单)是累计剩余的平方和. 最小二乘直线就是使剩余平方和最小的直线 $y = \beta_0 + \beta_1 x$. 由于数据的全部误差都假定在 y - 坐标之上，因此这条直线也称作 **y 在 x 上的回归直线**(line of regression of y on x). 直线的系数 β_0 和 β_1 称作(线性)**回归系**

1. 指代最小二乘直线时常采用这个记号，而不是 $y = mx + b$.

419 数(regression coefficients).[1]

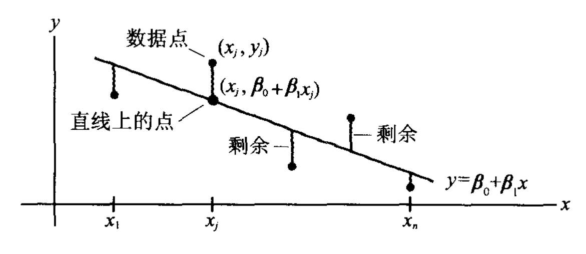


图6-27　用直线拟合实验数据

若数据点落在这条直线上，参数β_0和β_1将满足方程：

y的预测值		y的观测值
$\beta_0+\beta_1 x_1$	=	y_1
$\beta_0+\beta_1 x_2$	=	y_2
$\vdots$		$\vdots$
$\beta_0+\beta_1 x_n$	=	y_n

我们可以将这个方程组写作：

$$\boldsymbol{X\beta}=\boldsymbol{y},\text{其中 }\boldsymbol{X}=\begin{bmatrix}1 & x_1\\ 1 & x_2\\ \vdots & \vdots\\ 1 & x_n\end{bmatrix},\quad \boldsymbol{\beta}=\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix},\quad \boldsymbol{y}=\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{bmatrix}\tag{1}$$

当然，如果数据点不落在同一条直线上，就不存在参数β_0、β_1，使得$\boldsymbol{X\beta}$中的预测y-值与$\boldsymbol{y}$中的观测y-值相等，且$\boldsymbol{X\beta}=\boldsymbol{y}$无解. 这是一个$\boldsymbol{Ax}=\boldsymbol{b}$的最小二乘问题，只不过是采用了不同的记号而已！

向量$\boldsymbol{X\beta}$与$\boldsymbol{y}$距离的平方恰好等于剩余的平方和. 使这个和取值最小的$\boldsymbol{\beta}$也使得向量$\boldsymbol{X\beta}$与$\boldsymbol{y}$距离达到最小. 这样，求$\boldsymbol{X\beta}=\boldsymbol{y}$的最小二乘解等价于求确定图6-27最小二乘直线的$\boldsymbol{\beta}$.

【例题1】　求最小二乘直线方程$y=\beta_0+\beta_1 x$，使之最佳拟合数据点(2,1)，(5,2)，(7,3)和(8,3).

解：利用数据的x-坐标构造(1)中矩阵$\boldsymbol{X}$，利用y-坐标构造其中的向量$\boldsymbol{y}$：

$$\boldsymbol{X}=\begin{bmatrix}1 & 2\\ 1 & 5\\ 1 & 7\\ 1 & 8\end{bmatrix},\quad \boldsymbol{y}=\begin{bmatrix}1\\ 2\\ 3\\ 3\end{bmatrix}$$

---

1. 如果测量值误差在x而不在y之上，在绘出点并且计算回归直线以前，简单交换数据(x_j,y_j)的坐标即可. 如果两个坐标都允许有误差，那么你可以选择这样的直线，使点到直线正交(垂直)距离的平方和最小. 见7.5节基础练习.

使用新记号，对 $X\beta = y$ 的最小二乘解，构造正规方程：

$$X^T X\beta = X^T y$$

即，计算

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

420

$$X^T y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

正规方程为

$$\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

于是

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}^{-1} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 142 & -22 \\ -22 & 4 \end{bmatrix} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 24 \\ 30 \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$$

这样，如图 6-28 所示，最小二乘直线的方程为

$$y = \frac{2}{7} + \frac{5}{14}x$$

■

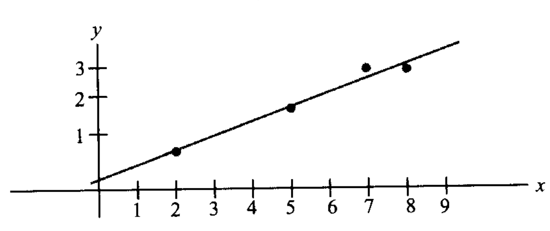


图 6-28 最小二乘直线 $y = \frac{2}{7} + \frac{5}{14}x$

在计算最小二乘直线之前，通常的做法是先计算原始 x - 值的平均值 $\bar{x}$，然后构造一个新变量 $x^* = x - \bar{x}$. 这个新的 x 值被称作是**平均偏差形式**(mean-deviation form). 此时，设计矩阵的两列互相正交. 正规方程的解将被简化，如 6.5 节例题 4. 详见习题 17 及 18.

6.6.2 一般线性模型

在一些实际应用中，数据点需要拟合成直线以外的其他模型. 后面的例题中矩阵方程仍是 $X\beta = y$，但不同问题中 X 的形式各不相同. 统计学家们通常引入**剩余向量**(residual vector) ϵ，定义成 $\epsilon = y - X\beta$，并且记

$$y = X\beta + \epsilon$$

具有上述形式的任意方程称作一个**线性模型**(linear model). 一旦 $\boldsymbol{X}$ 和 $\boldsymbol{y}$ 确定下来,

421 我们的目标就是最小化 $\boldsymbol{\epsilon}$ 的长度,相当于求 $\boldsymbol{X\beta}=\boldsymbol{y}$ 的最小二乘解. 无论何种情形,最小二乘解 $\hat{\boldsymbol{\beta}}$ 都是下列正规方程的解:

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X\beta}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}$$

6.6.3 其他曲线的最小二乘拟合

当散列图上的数据点 $(x_1,y_1),\cdots,(x_n,y_n)$ 不分布在任何直线附近时,可以假定 x 与 y 之间满足其他某种函数关系.

下面的三个例题演示了如何用具有一般形式

$$y=\beta_0f_0(x)+\beta_1f_1(x)+\cdots+\beta_kf_k(x) \tag{2}$$

的曲线来拟合数据,其中 $f_0,\cdots,f_k$ 是未知函数,$\beta_0,\cdots,\beta_k$ 是必须确定的参数. 正如我们将要看到的,由于方程(2)对未知参数来说是线性的,因此它描述了一个线性模型.

对于 x 的某个特定值,(2)给出了 y 的一个预测或"拟合"值. 观测值与预测值的差就是剩余. 参数 $\beta_0,\cdots,\beta_k$ 的选取必须使剩余的平方和达到最小.

【例题2】 假设数据点 $(x_1,y_1),\cdots,(x_n,y_n)$ 分布在某条抛物线而不是直线的附近. 例如,若 x 坐标记录某企业的生产水平,y 坐标记录在每天产出 x 单位的水平下单位产品的平均生产成本,则典型的平均成本曲线看起来是一条开口向上的抛物线(图6-29). 在生态学中,开口向下的抛物线通常用于描述植物生成的主要营养物质作为叶片表面积函数的模型(图6-30). 假如我们希望用形如

$$y=\beta_0+\beta_1x+\beta_2x^2 \tag{3}$$

的方程逼近这些数据. 试用方程(3)描述导出数据"最小二乘拟合"的线性模型.

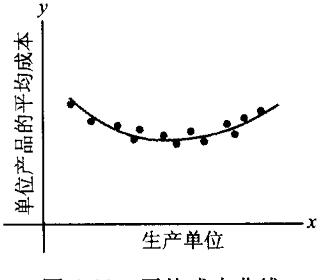


图6-29 平均成本曲线

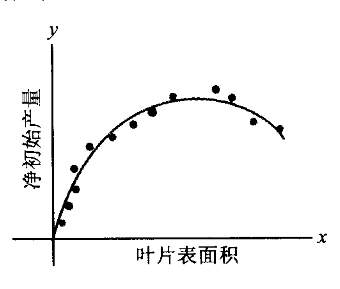


图6-30 营养物质的生产

解:方程(3)描述了一种理想的关系. 假设参数的实际取值为 β_0,β_1,β_2. 则第一个数据点的坐标满足下述形式的一个方程

$$y_1=\beta_0+\beta_1x_1+\beta_2x_1^2+\epsilon_1$$

其中 ϵ_1 是观测值 y_1 与预测 y-值 $\beta_0+\beta_1x_1+\beta_2x_1^2$ 之间的剩余误差. 每一个数据点都确定一个类似的方程

$$\begin{aligned}y_1&=\beta_0+\beta_1x_1+\beta_2x_1^2+\epsilon_1\\y_2&=\beta_0+\beta_1x_2+\beta_2x_2^2+\epsilon_2\\&\vdots\qquad\qquad\vdots\\y_n&=\beta_0+\beta_1x_n+\beta_2x_n^2+\epsilon_n\end{aligned}$$

422

易将这个方程组写成 $\boldsymbol{y}=\boldsymbol{X\beta}+\boldsymbol{\epsilon}$ 的形式. 为求出 $\boldsymbol{X}$，观察方程组的前几行，找出其中的规律.

$$\underset{\boldsymbol{y}}{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}} = \underset{\boldsymbol{X}}{\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix}} \underset{\boldsymbol{\beta}}{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}} + \underset{\boldsymbol{\epsilon}}{\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}}$$

■

【例题 3】 若数据点符合图 6-31 所示的模式，则一个适当的模型是具有形式

$$y=\beta_0+\beta_1x+\beta_2x^2+\beta_3x^3$$

的方程. 例如，将企业的总成本看作生产水平的函数，就是这样的数据. 试描述给出数据 $(x_1,y_1),\cdots,(x_n,y_n)$ 的这种最小二乘拟合的线性模型.

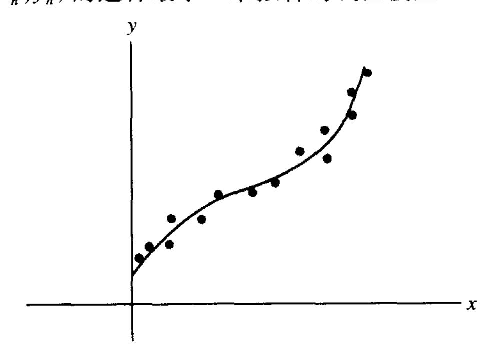


图 6-31 沿三次曲线分布的数据点

解：通过与例题 2 类似的分析，我们有：

观测向量　　设计矩阵　　参数向量　剩余向量

$$\boldsymbol{y}=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},\quad \boldsymbol{X}=\begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & x_n^3 \end{bmatrix},\quad \boldsymbol{\beta}=\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix},\quad \boldsymbol{\epsilon}=\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

■

6.6.4 多重回归

假设实验涉及两个独立的变量，比如 u 和 v，以及一个非独立变量 y. 用 u 和 v 预测 y 的一个简单方程形如

$$y=\beta_0+\beta_1u+\beta_2v \tag{4}$$

更一般的预测方程形如

$$y=\beta_0+\beta_1u+\beta_2v+\beta_3u^2+\beta_4uv+\beta_5v^2 \tag{5}$$

这个方程在地质学中经常用到，如，模拟地质侵蚀面、冰河谷、土壤 PH 值及其他量. 在这些情况下，最小二乘拟合称作趋势面.

(4)和(5)对未知参量来说是线性的(尽管 u 和 v 做了乘法)，因此它们各自导出一个线性模型. 一般地，使用形如

$$y=\beta_0 f_0(u,v)+\beta_1 f_1(u,v)+\cdots+\beta_k f_k(u,v)$$

423 的方程预测 y 时都能诱导一个线性模型，其中 $f_0,\cdots,f_k$ 是任意已知函数，$\beta_0,\cdots,\beta_k$ 是未知系数.

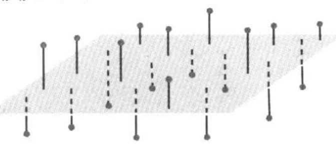


图 6-32 最小二乘平面

【例题 4】 地理学中，地形的局部模型由数据 $(u_1,v_1,y_1),\cdots,(u_n,v_n,y_n)$ 来构造，其中 u_j，v_j 和 y_j 分别为纬度、经度和海拔. 试描述基于(4)的给出数据最小二乘拟合的线性模型. 这个解称为最小二乘平面，见图 6-32.

解： 我们希望数据满足如下方程：

$$\begin{aligned}y_1&=\beta_0+\beta_1u_1+\beta_2v_1+\epsilon_1\\ y_2&=\beta_0+\beta_1u_2+\beta_2v_2+\epsilon_2\\ &\vdots\\ y_n&=\beta_0+\beta_1u_n+\beta_2v_n+\epsilon_n\end{aligned}$$

这个方程组有矩阵形式 $\boldsymbol{y}=\boldsymbol{X\beta}+\boldsymbol{\epsilon}$，其中

观测向量 设计矩阵 参数向量 剩余向量

$$\boldsymbol{y}=\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix},\quad \boldsymbol{X}=\begin{bmatrix}1&u_1&v_1\\1&u_2&v_2\\\vdots&\vdots&\vdots\\1&u_n&v_n\end{bmatrix},\quad \boldsymbol{\beta}=\begin{bmatrix}\beta_0\\\beta_1\\\beta_2\end{bmatrix},\quad \boldsymbol{\epsilon}=\begin{bmatrix}\epsilon_1\\\epsilon_2\\\vdots\\\epsilon_n\end{bmatrix}$$

■

例题 4 表明多重回归线性模型与先前例题中简单回归模型有相同的抽象形式. 线性代数使我们能够理解全体线性模型的一般原理. 一旦 $\boldsymbol{X}$ 被合理地定义，$\boldsymbol{\beta}$ 的正规方程不论涉及多少个变量，都具有相同的矩阵形式. 这样，对任意线性模型，若 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 可逆，最小二乘方 $\hat{\boldsymbol{\beta}}$ 由 $(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}$ 给出.

6.6.5 更多阅读

Ferguson, J., *Introduction to Linear Algebra in Geology* (New York: Chapman & Hall, 1994).

Krumbein, W. C., and F. A. Graybill, *An Introduction to Statistical Models in Geology* (New York: McGraw-Hill, 1965).

Legendre, P., and L. Legendre, *Numerical Ecology* (Amsterdam: Elsevier, 1998).

Unwin, David J., *An Introduction to Trend Surface Analysis*, Concepts and Techniques in 424 Modern Geography, No. 5 (Norwich, England: Geo Books, 1975).

基础练习

当产品的月销售服从季节性波动时，逼近销售数据的曲线可能形如

$$y=\beta_0+\beta_1x+\beta_2\sin(2\pi x/12)$$

其中 x 代表以月计的时间. 项 $\beta_0+\beta_1x$ 给出了基本的销售趋势，正弦项反映了销售的季节性变动. 给出诱导上述方程最小二乘拟合的线性模型的设计矩阵及参数向量. 假定数据为

$(x_1, y_1), \cdots, (x_n, y_n)$.

习题 6.6

在习题 1 ~4 中，求最佳拟合给定数据点的最小二乘直线方程 $y = \beta_0 + \beta_1 x$.

1. (0,1), (1,1), (2,2), (3,2) 2. (1,0), (2,1), (4,2), (5,3)
3. (−1,0), (0,1), (1,2), (2,4) 4. (2,3), (3,2), (5,1), (6,0)
5. 设 $\boldsymbol{X}$ 是求拟合数据 $(x_1, y_1), \cdots, (x_n, y_n)$ 的最小二乘直线时所用的设计矩阵. 利用 6.5 节的一个定理证明：正规方程有唯一的解当且仅当数据中包括至少两个 x-坐标不等的数据点.
6. 设 $\boldsymbol{X}$ 是例题 2 中的设计矩阵，对应于数据 $(x_1, y_1), \cdots, (x_n, y_n)$ 的最小二乘抛物线拟合. 假设 x_1, x_2 和 x_3 互不相等. 试从最小二乘方的角度说明为什么只有一条抛物线能最佳地拟合数据.
7. 某实验得到数据 (1,1.8), (2,2.7), (3,3.4), (4,3.8), (5,3.9). 试描述用形如
$$y = \beta_1 x + \beta_2 x^2$$
的函数最小二乘拟合这些点的模型. 例如，当产品销售供应量影响产品定价时，销售 x 单位产品所得的税收就属于这类函数.
   a. 给出设计矩阵、观测向量和未知的参数向量.
   b. [M]求出与数据相关的最小二乘曲线.
8. 销售水平 x 的函数 $y = \beta_1 x + \beta_2 x^2 + \beta_3 x^3$ 所对应的简单曲线，往往能为企业的各种成本提供好的模型. 函数中没有常数项是因为固定成本未计算在内.
   a. 利用数据 $(x_1, y_1), \cdots, (x_n, y_n)$，给出诱导上述方程最小二乘拟合的线性模型的设计矩阵和参数向量.
   b. [M]求具有上述形式的最小二乘曲线，使之拟合数据 (4,1.58), (6,2.08), (8,2.5), (10,2.8), (12,3.1), (14,3.4), (16,3.8), (18,4.32), 其中数值单位为千. 如果可能，绘出数据点以及三次逼近的图像.
9. 某实验得到数据 (1,7.9), (2,5.4) 和 (3, −0.9). 试描述用形如
$$y = A\cos x + B\sin x$$
的函数最小二乘拟合这些点的模型.
10. 设放射性物质 A 和 B 的裂变常数分别为 0.02 和 0.07. 若两种物质的一份混合物在 $t=0$ 时刻含有 M_A 克 A 和 M_B 克 B，则混合物在 t 时刻的总质量 y 有模型：
$$y = M_A e^{-0.02t} + M_B e^{-0.07t} \tag{6}$$
设初始值 M_A 和 M_B 未知，但科学家测得若干时刻下混合物的总质量，记为如下数据 (t_i, y_i): (10, 21.34), (11, 20.68), (12, 20.05), (14, 18.87), (15, 18.30).
   a. 试描述可用于估计 M_A 和 M_B 的线性模型.
   b. [M] 求基于(6)的最小二乘曲线.

425

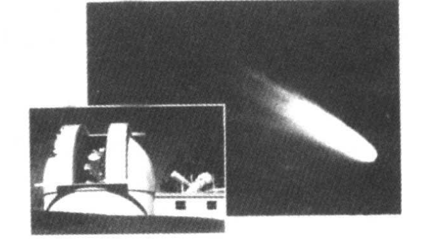


图 6-33 上次出现于 1986 年的哈雷彗星将在 2061 年重现

11. 根据开普勒第一定律，彗星拥有椭圆形、抛物线形或双曲线形的轨道(忽略来自各行星的万有引力). 在适当的极坐标下，彗星的位置 (r, ϑ) 满足下述形式的方程：
$$r = \beta + e(r \cdot \cos\vartheta)$$
其中 β 是常数，e 是轨道的离心率，且对椭圆有 $0 \leqslant e \leqslant 1$，对抛物线有 $e=1$，对双曲线有 $e>1$. 假设对新发现的彗星观测到如下数据.

确定轨道的类型，并预测$\vartheta=4.6$（弧度）时彗星所在的位置.[1]

ϑ	0.88	1.10	1.42	1.77	2.14
r	3.00	2.30	1.65	1.25	1.01

12. [**M**] 健康儿童的血压p(单位：毫米汞柱)和体重w(单位：磅)有近似下述方程的关系：

$$\beta_0+\beta_1 \ln w=p$$

利用下列实验数据估计体重100磅的健康儿童的血压：

w	44	61	81	113	131
$\ln w$	3.78	4.11	4.41	4.73	4.88
p	91	98	103	110	112

13. [**M**] 为测量飞机的起飞性能，从$t=0$到$t=12$以秒为间隔测量飞机的水平位置. 其位置(单位:英尺)为：0，8.8，29.9，62.0，104.7，159.1，222.0，294.5，380.4，471.1，571.7，686.8，809.2.

    a. 求数据的最小二乘三次曲线$y=\beta_0+\beta_1 t+\beta_2 t^2+\beta_3 t^3$.

    b. 利用(a)的结果估计$t=4.5$秒时飞机的速度.

14. 令$\bar{x}=\frac{1}{n}(x_1+\cdots+x_n)$，$\bar{y}=\frac{1}{n}(y_1+\cdots+y_n)$. 证明数据$(x_1,y_1),\cdots,(x_n,y_n)$的最小二乘直线一定通过点$(\bar{x},\bar{y})$. 即，证明$\bar{x}$和$\bar{y}$满足直线方程$\bar{y}=\hat{\beta}_0+\hat{\beta}_1\bar{x}$. [提示：这个方程由向量方程$\boldsymbol{y}=X\hat{\boldsymbol{\beta}}+\boldsymbol{\epsilon}$而得. 用**1**记$X$的首列. 利用剩余向量与$X$列空间正交从而与**1**正交的事实.]

给定最小二乘问题的数据$(x_1,y_1),\cdots,(x_n,y_n)$，下列缩写十分有用：

$$\sum x=\sum_{i=1}^n x_i,\ \sum x^2=\sum_{i=1}^n x_i^2,$$
$$\sum y=\sum_{i=1}^n y_i,\ \sum xy=\sum_{i=1}^n x_i y_i$$

最小二乘直线$y=\hat{\beta}_0+\hat{\beta}_1 x$的正规方程可以写成如下形式：

$$n\hat{\beta}_0+\hat{\beta}_1\sum x=\sum y$$
$$\hat{\beta}_0\sum x+\hat{\beta}_1\sum x^2=\sum xy \qquad (7)$$

15. 从本节给出的矩阵形式出发，推导正规方程(7).
16. 利用矩阵的逆解方程组(7)，进而给出统计学教材中常见的$\hat{\beta}_0$和$\hat{\beta}_1$的公式.
17. a. 将例题1中数据重写为新x-坐标的平均偏差形式. 令X是对应的设计矩阵. X的各列为何正交?

    b. 写出(a)中数据的正规方程，求解这些方程并求最小二乘直线$\boldsymbol{y}=\beta_0+\beta_1\boldsymbol{x}^*$，其中$\boldsymbol{x}^*=\boldsymbol{x}-5.5$.

426

18. 设数据$(x_1,y_1),\cdots,(x_n,y_n)$的$x$坐标为平均偏差形式，且满足$\sum x_i=0$. 证明：若$X$是最小二乘直线的设计矩阵，则$X^TX$是对角阵.

习题19和20涉及含有两列或多列的设计矩阵X以及$\boldsymbol{y}=X\boldsymbol{\beta}$的最小二乘解$\hat{\boldsymbol{\beta}}$. 考虑下列各数：

---

1. 最小二乘拟合的基本思想归功于高斯(同时也是A. 勒让德的独立工作的成果)，1801年高斯使用这一方法确定了小行星谷神星的轨道，一举成名. 在发现这颗行星的40天后，它消失在太阳的背面. 高斯预言它十个月之后会重新出现并给出了它的位置. 预言的准确性震惊了欧洲科学界.

(i) $\|X\hat{\boldsymbol{\beta}}\|^2$——“回归项”的平方和，记为SS(R).

(ii) $\|\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\|^2$——误差项的平方和，记为SS(E).

(iii) $\|\boldsymbol{y}\|^2$——y值的平方总和，记为SS(T).

所有的统计学教材，凡是讨论了回归和线性模型$\boldsymbol{y}=X\boldsymbol{\beta}+\epsilon$，都会介绍这些值，尽管采用的术语和记号可能不尽相同. 简单起见，假设y-值的平均值为零. 此时，SS(T)与y-值集的方差成正比.

19. 验证方程SS(T) = SS(R) + SS(E). [提示：利用一个定理，并且说明定理的假设为何被满足.]这个方程在统计学的回归理论和方差分析中都极为重要.
20. 证明$\|X\hat{\boldsymbol{\beta}}\|^2=\hat{\boldsymbol{\beta}}^{\mathrm{T}}X^{\mathrm{T}}\boldsymbol{y}$. [提示：重写方程左端，并利用$\hat{\boldsymbol{\beta}}$满足正规方程的事实.]SS(R)的这个公式在统计学中十分有用. 根据这一公式以及习题19，得到SS(E)的标准公式：

$$\mathrm{SS(E)}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}-\hat{\boldsymbol{\beta}}^{\mathrm{T}}X^{\mathrm{T}}\boldsymbol{y}$$

基础练习答案

构造X和$\boldsymbol{\beta}$，使得$X\boldsymbol{\beta}$的第k行是与数据点(x_k,y_k)对应的预测y值，即

$$\beta_0+\beta_1x_k+\beta_2\sin(2\pi x_k/12)$$

显然有

$$X=\begin{bmatrix}1 & x_1 & \sin(2\pi x_1/12)\\ \vdots & \vdots & \vdots\\ 1 & x_n & \sin(2\pi x_n/12)\end{bmatrix},\quad \boldsymbol{\beta}=\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\end{bmatrix}$$

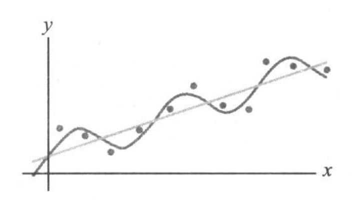


图6-34 呈现季节性波动的销售量

6.7 内积空间

在含有向量空间的实际应用中，长度、距离和正交性等概念通常都很重要. 对$\mathbf{R}^n$来说，这些概念以6.1节定理1所罗列的内积性质为基础. 对其他空间，我们需要类推具有同样性质的内积. 这样，定理1的结论就是下面定义中的公理：

427

【定义】 向量空间V上的一个**内积**(inner product)是满足如下性质的一个函数：V中的每个向量对$\boldsymbol{u}$，$\boldsymbol{v}$，都对应一个实数值$\langle\boldsymbol{u},\boldsymbol{v}\rangle$，且$V$中的任意$\boldsymbol{u}$，$\boldsymbol{v}$和$\boldsymbol{w}$及任意数量$c$都满足下列公理：

1. $\langle\boldsymbol{u},\boldsymbol{v}\rangle=\langle\boldsymbol{v},\boldsymbol{u}\rangle$
2. $\langle\boldsymbol{u}+\boldsymbol{v},\boldsymbol{w}\rangle=\langle\boldsymbol{u},\boldsymbol{w}\rangle+\langle\boldsymbol{v},\boldsymbol{w}\rangle$
3. $\langle c\boldsymbol{u},\boldsymbol{v}\rangle=c\langle\boldsymbol{u},\boldsymbol{v}\rangle$
4. $\langle\boldsymbol{u},\boldsymbol{u}\rangle\geqslant 0$，且$\langle\boldsymbol{u},\boldsymbol{u}\rangle=0$当且仅当$\boldsymbol{u}=\boldsymbol{0}$

赋予了内积的向量空间，称作**内积空间**(inner product space).

具有标准内积的向量空间$\mathbf{R}^n$是一个内积空间，并且本章中对$\mathbf{R}^n$讨论的几乎全部结论都可推广到一般内积空间. 本节以及下一节的例题将为工程学、物理学、数学和统计学等学科中的各种实际应用奠定基础.

【例题1】 固定两个正数，比如4和5，对$\mathbf{R}^2$中向量$\boldsymbol{u}=(u_1,u_2)$以及$\boldsymbol{v}=(v_1,v_2)$，令

$$\langle\boldsymbol{u},\boldsymbol{v}\rangle=4u_1v_1+5u_2v_2 \tag{1}$$

证明(1)定义了一个内积.

解： 由于$\langle \boldsymbol{u},\boldsymbol{v}\rangle = 4u_1v_1+5u_2v_2=4v_1u_1+5v_2u_2=\langle \boldsymbol{v},\boldsymbol{u}\rangle$，公理1自然被满足. 若$\boldsymbol{w}=(w_1,w_2)$，则

$$\begin{aligned}\langle \boldsymbol{u}+\boldsymbol{v},\boldsymbol{w}\rangle &= 4(u_1+v_1)w_1+5(u_2+v_2)w_2\\ &= 4u_1w_1+5u_2w_2+4v_1w_1+5v_2w_2\\ &= \langle \boldsymbol{u},\boldsymbol{w}\rangle+\langle \boldsymbol{v},\boldsymbol{w}\rangle\end{aligned}$$

这就验证了公理2. 对于公理3，我们有

$$\langle c\boldsymbol{u},\boldsymbol{v}\rangle = 4(cu_1)v_1+5(cu_2)v_2=c(4u_1v_1+5u_2v_2)=c\langle \boldsymbol{u},\boldsymbol{v}\rangle$$

对于公理4，注意到$\langle \boldsymbol{u},\boldsymbol{u}\rangle=4u_1^2+5u_2^2\geqslant 0$，并且$4u_1^2+5u_2^2=0$当且仅当$u_1=u_2=0$，即$\boldsymbol{u}=\boldsymbol{0}$. 此外还有$\langle \boldsymbol{0},\boldsymbol{0}\rangle=0$. 所以(1)定义了$\mathbf{R}^2$上的一个内积. ■

$\mathbf{R}^n$上也可以定义与(1)相似的内积. 它们自然地和“加权最小二乘”问题联系起来，其中赋给内积和式各个分量的权值满足：测量数据越可靠，所赋予的权值越大.

从现在起，当内积空间涉及多项式或者其他函数时，我们将不对向量使用粗体字体，而按照惯常方式来书写. 不过，必须记住，当把函数看成是向量空间的元素时，每个函数都是一个向量.

428

【例题2】 令$t_0,\cdots,t_n$是互不相同的实数. 对于$\mathbf{P}_n$中的p，q，定义

$$\langle p,q\rangle = p(t_0)q(t_0)+p(t_1)q(t_1)+\cdots+p(t_n)q(t_n) \tag{2}$$

内积公理1~3容易验证. 对于公理4，注意

$$\langle p,p\rangle = [p(t_0)]^2+[p(t_1)]^2+\cdots+[p(t_n)]^2\geqslant 0$$

另外还有$\langle \boldsymbol{0},\boldsymbol{0}\rangle=0$(粗体的零代表零多项式，即$\mathbf{P}_n$的零向量). 若$\langle p,p\rangle=0$，则$p$必须在$t_0,\cdots,t_n$这$n+1$个点上取零值. 这仅在$p$是零多项式时才有可能，因为$p$的次数小于$n+1$. 于是(2)定义了$\mathbf{P}_n$上的一个内积. ■

【例题3】 设V是赋予了例题2所述内积的$\mathbf{P}_2$空间，其中$t_0=0$，$t_1=\frac{1}{2}$，$t_2=1$. 令$p(t)=12t^2$，$q(t)=2t-1$. 计算$\langle p,q\rangle$和$\langle q,q\rangle$.

解：

$$\begin{aligned}\langle p,q\rangle &= p(0)q(0)+p(\tfrac{1}{2})q(\tfrac{1}{2})+p(1)q(1)\\ &= (0)(-1)+(3)(0)+(12)(1)=12\\ \langle q,q\rangle &= [q(0)]^2+\left[q(\tfrac{1}{2})\right]^2+[q(1)]^2\\ &= (-1)^2+(0)^2+(1)^2=2\end{aligned}$$ ■

6.7.1　长度、距离和正交性

设V是一个内积空间，其内积记号为$\langle \boldsymbol{u},\boldsymbol{v}\rangle$. 类似于$\mathbf{R}^n$，我们将向量$\boldsymbol{v}$的**长度**(length)或**范式**(norm)定义为下面的数量：

$$\|\boldsymbol{v}\| = \sqrt{\langle \boldsymbol{v},\boldsymbol{v}\rangle}$$

或者等价地，$\|\boldsymbol{v}\|^2=\langle \boldsymbol{v},\boldsymbol{v}\rangle$. (由于$\langle \boldsymbol{v},\boldsymbol{v}\rangle\geqslant 0$，定义有意义. 但这个定义并不能说明$\langle \boldsymbol{v},\boldsymbol{v}\rangle$是“平方和”，这是因为$\boldsymbol{v}$不一定是$\mathbf{R}^n$中的元素.)

一个**单位向量**(unit vector)是长度为1的向量. $\boldsymbol{u}$**到**$\boldsymbol{v}$**的距离**(distance between $\boldsymbol{u}$

and $\boldsymbol{v}$)是$\|\boldsymbol{u}-\boldsymbol{v}\|$. 向量$\boldsymbol{u}$和$\boldsymbol{v}$**正交**(orthogonal)，如果$\langle \boldsymbol{u},\boldsymbol{v}\rangle=0$.

【例题4】 设$\mathbf{P}_2$具有例题2中(2)所述的内积. 计算向量$p(t)=12t^2$和$q(t)=2t-1$的长度.

解：

$$\|p\|^2=\langle p,p\rangle=[p(0)]^2+\left[p(\frac{1}{2})\right]^2+[p(1)]^2$$

$$=0+[3]^2+[12]^2=153$$

$$\|p\|=\sqrt{153}$$

由例题3，$\langle q,q\rangle=2$. 于是$\|\boldsymbol{q}\|=\sqrt{2}$. ■ 429

6.7.2 格拉姆-施密特方法

与$\mathbf{R}^n$相同，对于内积空间的有限维子空间，可以用格拉姆-施密特方法来确定其正交基的存在性. 使用这一方法可以构造出实际应用中常用的一些正交基.

向量在含正交基的子空间W上的正交投影也可以用通常方法加以构造. 这个投影不依赖于正交基的选取，并且具有正交分解定理、最佳逼近定理所述的性质.

【例题5】 设V是赋予了例题2所述内积的$\mathbf{P}_4$空间，其中多项式在-2，-1，0，1和2各点上有赋值. 将$\mathbf{P}_2$看作V的一个子空间. 对多项式1，t和t^2应用格拉姆-施密特方法，求$\mathbf{P}_2$的一组正交基.

解： 由于内积只依赖于多项式在$-2,\cdots,2$处的赋值，因此我们将多项式在这些点的值看作$\mathbf{R}^5$中的向量，列在多项式下方：[1]

$$\text{多项式：}\quad 1 \qquad t \qquad t^2$$

$$\text{值向量：}\begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix},\begin{bmatrix}-2\\-1\\0\\1\\2\end{bmatrix},\begin{bmatrix}4\\1\\0\\1\\4\end{bmatrix}$$

V中两多项式的内积等于它们在$\mathbf{R}^5$中对应向量的(标准)内积. 注意到t与常函数1正交. 故取$p_0(t)=1$，$p_1(t)=t$. 对于p_2，用$\mathbf{R}^5$中的向量计算t^2在$\mathrm{Span}\{p_0,p_1\}$上的投影：

$$\langle t^2,p_0\rangle=\langle t^2,1\rangle=4+1+0+1+4=10$$

$$\langle p_0,p_0\rangle=5$$

$$\langle t^2,p_1\rangle=\langle t^2,t\rangle=-8+(-1)+0+1+8=0$$

t^2在$\mathrm{Span}\{1,t\}$上的正交投影是$\frac{10}{5}p_0+0p_1$. 于是

$$p_2(t)=t^2-2p_0(t)=t^2-2$$

430

V子空间$\mathbf{P}_2$的一组正交基为：

---

1. $\mathbf{P}_4$中的每一多项式由它在$-2,\cdots,2$这五个点处的值唯一确定. 事实上，从多项式p到其值向量的对应是同构，即，到$\mathbf{R}^5$且保持线性组合的一一映射.

$$\text{多项式:}\quad p_0 \qquad p_1 \qquad p_2$$

$$\text{值向量:}\quad \begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix},\begin{bmatrix}-2\\-1\\0\\1\\2\end{bmatrix},\begin{bmatrix}2\\-1\\-2\\-1\\2\end{bmatrix} \tag{3}$$

■

6.7.3 内积空间中的最佳逼近

应用数学中一个较为普遍的问题涉及以函数为元素的向量空间 V. 这类问题要求用 V 的指定子空间 W 中的函数 g 逼近 V 中的函数 f. f 的逼近是否"接近"取决于 $\|f-g\|$ 的定义. 我们只考虑 f 和 g 之间距离由内积定义的情形. 此时, W 中函数对 f 的最佳逼近就是 f 在子空间 W 上的正交投影.

【例题 6】　设 V 是赋予了例题 5 所述内积的 $\mathbf{P}_4$ 空间, 并且设 p_0, p_1 和 p_2 是例题 5 所求的子空间 $\mathbf{P}_2$ 的正交基. 试求 $\mathbf{P}_2$ 中多项式对 $p(t)=5-\frac{1}{2}t^4$ 的最佳逼近.

解: p_0, p_1 和 p_2 在 -2, -1, 0, 1 和 2 各处的赋值可以列成 $\mathbf{R}^5$ 中的向量, 如前文(3)所述. p 的对应赋值为 -3, 9/2, 5, 9/2 和 -3. 计算:

$$\langle p,p_0\rangle=8,\qquad \langle p,p_1\rangle=0,\qquad \langle p,p_2\rangle=-31$$
$$\langle p_0,p_0\rangle=5,\qquad\qquad \langle p_2,p_2\rangle=14$$

则 $\mathbf{P}_2$ 中多项式对 p 的最佳逼近是:

$$\begin{aligned}\hat{p}&=\operatorname{proj}_{\mathbf{P}_2}p=\frac{\langle p,p_0\rangle}{\langle p_0,p_0\rangle}p_0+\frac{\langle p,p_1\rangle}{\langle p_1,p_1\rangle}p_1+\frac{\langle p,p_2\rangle}{\langle p_2,p_2\rangle}p_2\\&=\frac{8}{5}p_0+\frac{-31}{14}p_2=\frac{8}{5}-\frac{31}{14}(t^2-2)\end{aligned}$$

当多项式之间的距离只用 -2, -1, 0, 1 和 2 各点的赋值来度量时, 上述多项式是 $\mathbf{P}_2$ 中与 p 距离最近的多项式. 见图 6-35. ■

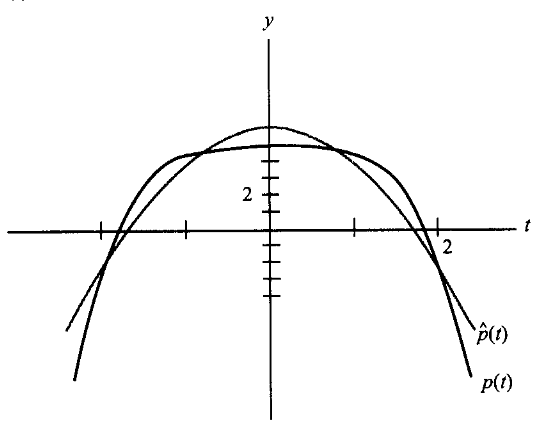


图　6-35

例题5和6中的p_0，p_1和p_2所属的一类多项式，在统计学中称作正交多项式.[1]这种正交性归属于例题2所述的那类内积.

431

6.7.4 两个不等式

给定内积空间V的一个向量$\boldsymbol{v}$和一个有限维子空间W，我们可以对$\boldsymbol{v}$关于W的正交分解应用勾股定理，并得到

$$\|\boldsymbol{v}\|^2 = \|\mathrm{proj}_W\boldsymbol{v}\|^2 + \|\boldsymbol{v} - \mathrm{proj}_W\boldsymbol{v}\|^2$$

如图6-36. 特别地，上式表明$\boldsymbol{v}$在W上投影的范式不会超过$\boldsymbol{v}$本身的范式. 这一简单事实可以导出下面重要的不等式.

【定理16】 柯西－施瓦兹不等式

对V中任意$\boldsymbol{u}$，$\boldsymbol{v}$，

$$|\langle \boldsymbol{u},\boldsymbol{v}\rangle| \leqslant \|\boldsymbol{u}\|\|\boldsymbol{v}\| \tag{4}$$

证明：若$\boldsymbol{u}=\boldsymbol{0}$，则(4)式的两端均为零，于是此时(4)式成立(见基础练习1). 若$\boldsymbol{u}\neq\boldsymbol{0}$，令$W$为由$\boldsymbol{u}$张成的子空间. 回忆：对任意数量$c$，都有$\|c\boldsymbol{u}\| = |c|\|\boldsymbol{u}\|$. 于是，

$$\|\mathrm{proj}_W\boldsymbol{v}\| = \left\|\frac{\langle \boldsymbol{v},\boldsymbol{u}\rangle}{\langle \boldsymbol{u},\boldsymbol{u}\rangle}\boldsymbol{u}\right\| = \frac{|\langle \boldsymbol{v},\boldsymbol{u}\rangle|}{|\langle \boldsymbol{u},\boldsymbol{u}\rangle|}\|\boldsymbol{u}\| = \frac{|\langle \boldsymbol{v},\boldsymbol{u}\rangle|}{\|\boldsymbol{u}\|^2}\|\boldsymbol{u}\| = \frac{|\langle \boldsymbol{u},\boldsymbol{v}\rangle|}{\|\boldsymbol{u}\|}$$

由于$\|\mathrm{proj}_W\boldsymbol{v}\| \leqslant \|\boldsymbol{v}\|$，我们有$\dfrac{|\langle \boldsymbol{u},\ \boldsymbol{v}\rangle|}{\|\boldsymbol{u}\|} \leqslant \|\boldsymbol{v}\|$，进而可得(4). ■

柯西－施瓦兹不等式在许多数学分支中都非常有用，习题中给出了它的几个简单应用. 我们这里主要使用它来证明另一个涉及向量范式的基本不等式，见图6-37.

432

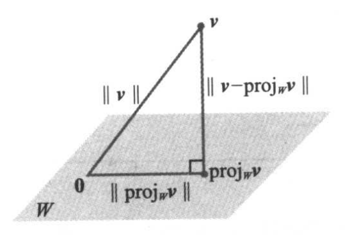


图6-36 直角三角形的斜边为最长边

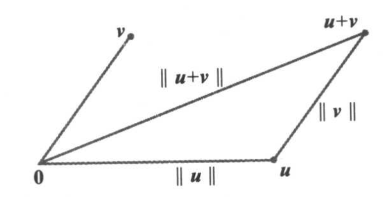


图6-37 三角形各边长度

【定理17】 三角不等式

对V中任意$\boldsymbol{u}$，$\boldsymbol{v}$，

$$\|\boldsymbol{u}+\boldsymbol{v}\| \leqslant \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$$

---

1. 见Norman L. Johnson and Fred C. Leone，*Statistics and Experimental Design in Engineering and the Physical Sciences*（New York：John Wiley & Sons，1964），pp. 424－436. 该书第430－431页上的表格列出了“正交多项式”，实际就是多项式在某些点(比如－2，－1，0，1和2)上的赋值.

证明：

$$\begin{aligned}\|u+v\|^2 &= \langle u+v,u+v\rangle = \langle u,u\rangle + 2\langle u,v\rangle + \langle v,v\rangle \\ &\leqslant \|u\|^2 + 2|\langle u,v\rangle| + \|v\|^2 \\ &\leqslant \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \qquad \text{柯西－施瓦兹不等式} \\ &= (\|u\| + \|v\|)^2\end{aligned}$$

上式两端开平方根，即得三角不等式. ■

6.7.5　$C[a,b]$的一个内积（需要微积分知识）

区间$a\leqslant t\leqslant b$上全体连续函数所构成的向量空间$C[a,b]$可能是实际应用中使用最广泛的内积空间，其上的内积将在下文加以介绍.

我们首先考虑一个多项式p和一个整数n，这里n大于或等于p的次数. 则p属于$\mathbf{P}_n$，并且，采用例题2中由$[a,b]$区间内$n+1$个点处赋值来定义的内积，我们可以计算p的“长度”. 然而，这个长度只提取了多项式在这$n+1$个点处的性质. 由于对于更大的n，p都属于$\mathbf{P}_n$，因此我们可以选取更大的n，利用更多点处的赋值计算对应的“赋值”内积，见图6-38.

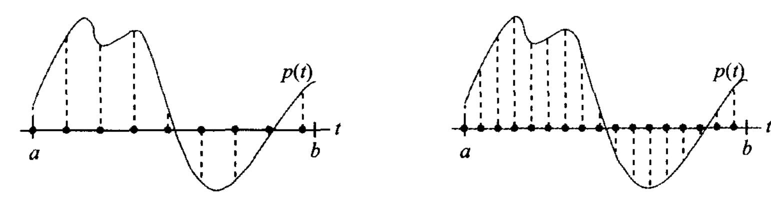
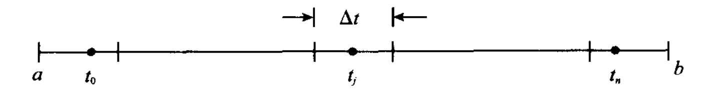


图6-38　利用$[a,b]$内不同数目点的赋值计算$\|p\|^2$

我们将$[a,b]$划分成$n+1$个长为$\Delta t=(b-a)/(n+1)$的子区间，并且令$t_0,\cdots,t_n$为这些子区间内的任意一点.



如果n充分大，$\mathbf{P}_n$上由$t_0,\cdots,t_n$确定的内积很可能使$\langle p,p\rangle$取值非常大，所以我们将它除以$n+1$以缩小这个值. 注意到$1/(n+1)=\Delta t/(b-a)$，定义：

433

$$\langle p,q\rangle = \frac{1}{n+1}\sum_{j=0}^{n} p(t_j)q(t_j) = \frac{1}{b-a}\left[\sum_{j=0}^{n} p(t_j)q(t_j)\Delta t\right]$$

现在让n无限增大. 因为多项式p，q是连续函数，方括号内的表达式是一个趋向某定积分的黎曼和，此外我们考虑$[a,b]$区间上$p(t)q(t)$的平均值：

$$\frac{1}{b-a}\int_a^b p(t)q(t)\,\mathrm{d}t$$

这个量对任意次数的多项式都有定义（事实上对全体连续函数均可），并且如下面的例题所言，它具有内积的所有性质. 缩放因子$1/(b-a)$不是必须的，简单起见常将它省略.

【例题7】　对$C[a,\ b]$中的f和g，令

$$\langle f,g\rangle = \int_a^b f(t)g(t)\,dt \tag{5}$$

证明(5)定义了 $C[a,b]$上的一个内积.

解：定积分的基本性质可以推出内积公理 1 ~3. 对于公理 4，注意到：

$$\langle f,f\rangle = \int_a^b [f(t)]^2 dt \geqslant 0$$

函数$[f(t)]^2$连续且在$[a,b]$上非负. 若$[f(t)]^2$的定积分为零，则由高等微积分的一个定理可知$[f(t)]^2$在$[a,b]$上恒等于零，此时f必为零函数. 这样，$\langle f,f\rangle=0$ 蕴含f是$[a,b]$上的零函数. 因此(5)定义 $C[a,b]$上的一个内积. ■

【例题 8】 设 V 是赋予了例题7 所述内积的 $C[0,1]$空间，令 W 是多项式 $p_1(t)=1$，$p_2(t)=2t-1$和 $p_3(t)=12t^2$张成的子空间. 试用格拉姆 - 施密特方法求 W 的一组正交基.

解：令 $q_1=p_1$，计算

$$\langle p_2,q_1\rangle = \int_0^1 (2t-1)(1)\,dt = (t^2-t)\Big|_0^1 = 0$$

所以 p_2已经正交于 q_1，于是我们取 $q_2=p_2$. 下面计算 p_3在 $W_2=\text{Span}\{q_1,q_2\}$上的投影：

$$\langle p_3,q_1\rangle = \int_0^1 12t^2\cdot 1\,dt = 4t^3\Big|_0^1 = 4$$

434

$$\langle q_1,q_1\rangle = \int_0^1 1\cdot 1\,dt = t\Big|_0^1 = 1$$

$$\langle p_3,q_2\rangle = \int_0^1 12t^2(2t-1)\,dt = \int_0^1 (24t^3-12t^2)\,dt = 2$$

$$\langle q_2,q_2\rangle = \int_0^1 (2t-1)^2 dt = \frac{1}{6}(2t-1)^3\Big|_0^1 = \frac{1}{3}$$

于是

$$\text{proj}_{W_2}p_3 = \frac{\langle p_3,q_1\rangle}{\langle q_1,q_1\rangle}q_1 + \frac{\langle p_3,q_2\rangle}{\langle q_2,q_2\rangle}q_2 = \frac{4}{1}q_1 + \frac{2}{1/3}q_2 = 4q_1+6q_2$$

令

$$q_3 = p_3 - \text{proj}_{W_2}p_3 = p_3 - 4q_1 - 6q_2$$

写成函数有，$q_3(t)=12t^2-4-6(2t-1)=12t^2-12t+2$. 求得子空间 W 的一组正交基为$\{q_1,q_2,q_3\}$. ■

基础练习

利用内积公理验证下列论断：

1. $\langle \boldsymbol{v},\boldsymbol{0}\rangle=\langle \boldsymbol{0},\boldsymbol{v}\rangle=0$.
2. $\langle \boldsymbol{u},\boldsymbol{v}+\boldsymbol{w}\rangle=\langle \boldsymbol{u},\boldsymbol{v}\rangle+\langle \boldsymbol{u},\boldsymbol{w}\rangle$.

习题 6.7

1. 设 $\mathbf{R}^2$具有例题 1 所述的内积，并且令 $\boldsymbol{x}=(1,1)$，$\boldsymbol{y}=(5,-1)$.
   a. 求$\|\boldsymbol{x}\|$，$\|\boldsymbol{y}\|$和$|\langle \boldsymbol{x},\boldsymbol{y}\rangle|^2$.
   b. 写出与 $\boldsymbol{y}$ 正交的全体向量(z_1,z_2).

2. 设 $\mathbf{R}^2$ 具有例题 1 所述的内积，证明柯西－施瓦兹不等式对 $\boldsymbol{x}=(3,-2)$，$\boldsymbol{y}=(-2,1)$ 成立. [建议：考虑 $|\langle \boldsymbol{x},\boldsymbol{y}\rangle|^2$.]

习题 3～8 均在 $\mathbf{P}_2$ 内讨论，其内积由 -1，0 和 1 处的赋值确定.

3. 计算 $\langle p,q\rangle$，其中 $p(t)=4+t$，$q(t)=5-4t^2$.
4. 计算 $\langle p,q\rangle$，其中 $p(t)=3t-t^2$，$q(t)=3+2t^2$.
5. 对习题 3 中的 p 和 q，计算 $\|p\|$ 和 $\|q\|$.
6. 对习题 4 中的 p 和 q，计算 $\|p\|$ 和 $\|q\|$.
7. 对习题 3 中的 p 和 q，计算 q 在 p 所张成的子空间上的正交投影.
8. 对习题 4 中的 p 和 q，计算 q 在 p 所张成的子空间上的正交投影.
9. 设 $\mathbf{P}_3$ 上有内积，且由 -3，-1，1 和 3 处的赋值确定. 令 $p_0(t)=1$，$p_1(t)=t$ 且 $p_2(t)=t^2$.
   a. 计算 p_2 在 p_0,p_1 所张成的子空间上的正交投影.
   b. 求同时与 p_0，p_1 正交的一个多项式 q，使得 $\{p_0,p_1,q\}$ 构成 Span$\{p_0,p_1,p_2\}$ 的一组正交基. 适当缩放多项式 q，使其在 $(-3,-1,1,3)$ 处的值向量等于 $(1,-1,-1,1)$.
10. 设 $\mathbf{P}_3$ 具有习题 9 所述的内积，p_0,p_1 和 q 也如习题 9 所述. 试求 Span$\{p_0,p_1,q\}$ 中多项式对 $p(t)=t^3$ 的最佳逼近.
11. 设 p_0，p_1 和 p_2 是例题 5 中所述的正交多项式，其中 $\mathbf{P}_4$ 上的内积由 -2，-1，0，1 和 2 处的赋值所确定. 试求 t^3 在 Span$\{p_0,p_1,p_2\}$ 上的正交投影.
12. 求一多项式 p_3，使得 $\{p_0,p_1,p_2,p_3\}$(见习题 11)构成 $\mathbf{P}_4$ 子空间 $\mathbf{P}_3$ 的一组正交基. 适当缩放多项式，使其值向量等于 $(-1,2,0,-2,1)$.

435

13. 设 A 是任意可逆 $n\times n$ 矩阵. 证明对于 $\boldsymbol{u}$，$\boldsymbol{v}\in\mathbf{R}^n$，公式 $\langle \boldsymbol{u},\boldsymbol{v}\rangle=(A\boldsymbol{u})\cdot(A\boldsymbol{v})=(A\boldsymbol{u})^{\mathrm{T}}(A\boldsymbol{v})$ 定义了 $\mathbf{R}^n$ 上的一个内积.
14. 设 T 是从向量空间 V 到 $\mathbf{R}^n$ 的一个一对一的线性变换. 证明对于 $\boldsymbol{u}$，$\boldsymbol{v}\in V$，公式 $\langle \boldsymbol{u},\boldsymbol{v}\rangle=T(\boldsymbol{u})\cdot T(\boldsymbol{v})$ 定义了 V 上的一个内积.

利用内积公理及本节其他结论，验证习题 15～18 中的命题.

15. 对任意数量 c，$\langle \boldsymbol{u},c\boldsymbol{v}\rangle=c\langle \boldsymbol{u},\boldsymbol{v}\rangle$.
16. 如果 $\{\boldsymbol{u},\boldsymbol{v}\}$ 是 V 的一个标准正交集，则 $\|\boldsymbol{u}-\boldsymbol{v}\|=\sqrt{2}$.
17. $\langle \boldsymbol{u},\boldsymbol{v}\rangle=\frac{1}{4}\|\boldsymbol{u}+\boldsymbol{v}\|^2-\frac{1}{4}\|\boldsymbol{u}-\boldsymbol{v}\|^2$.
18. $\|\boldsymbol{u}+\boldsymbol{v}\|^2+\|\boldsymbol{u}-\boldsymbol{v}\|^2=2\|\boldsymbol{u}\|^2+2\|\boldsymbol{v}\|^2$.
19. 给定 $a\geqslant 0$ 和 $b\geqslant 0$，并且令 $\boldsymbol{u}=\begin{bmatrix}\sqrt{a}\\ \sqrt{b}\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}\sqrt{b}\\ \sqrt{a}\end{bmatrix}$. 利用柯西－施瓦兹不等式比较几何平均值 $\sqrt{ab}$ 与算术平均值 $(a+b)/2$.
20. 令 $\boldsymbol{u}=\begin{bmatrix}a\\ b\end{bmatrix}$，$\boldsymbol{v}=\begin{bmatrix}1\\ 1\end{bmatrix}$. 利用柯西－施瓦兹不等式证明：

$$\left(\frac{a+b}{2}\right)^2\leqslant\frac{a^2+b^2}{2}$$

习题 21～24 均在 $V=C[0,1]$ 内讨论，其内积由例题 7 中的积分给定.

21. 计算 $\langle f,g\rangle$，其中 $f(t)=1-3t^2$，$g(t)=t-t^3$.
22. 计算 $\langle f,g\rangle$，其中 $f(t)=5t-3$，$g(t)=t^3-t^2$.
23. 对习题 21 中的 f，计算 $\|f\|$.
24. 对习题 22 中的 g，计算 $\|g\|$.

25. 设 V 是具有例题7所述内积的 $C[-1,1]$ 空间. 试求由1, t 和 t^2 所张成子空间的一组正交基. 这组基中的多项式称作勒让德多项式.
26. 设 V 是具有例题7所述内积的 $C[-2,2]$ 空间. 试求由1, t 和 t^2 所张成子空间的一组正交基.
27. [M] 设 $\mathbf{P}_4$ 有例题5所述的内积, 并设 p_0, p_1, p_2 为该题中的正交多项式. 利用你的矩阵软件, 对集合 $\{p_0,p_1,p_2,t^3,t^4\}$ 应用格拉姆-施密特方法, 构造 $\mathbf{P}_4$ 的一组正交基.
28. [M] 设 V 是具有例题7所述内积的 $C[0,2\pi]$ 空间. 利用格拉姆-施密特方法, 构造 $\{1,\cos t,\cos^2 t,\cos^3 t\}$ 所张成子空间的一组正交基. 利用矩阵软件或计算程序计算相应的定积分.

基础练习答案

1. 由公理1可知 $\langle \boldsymbol{v},\boldsymbol{0}\rangle=\langle \boldsymbol{0},\boldsymbol{v}\rangle$. 又由公理3有, $\langle \boldsymbol{0},\boldsymbol{v}\rangle=\langle 0\boldsymbol{v},\boldsymbol{v}\rangle=0\langle \boldsymbol{v},\boldsymbol{v}\rangle$, 故 $\langle \boldsymbol{0},\boldsymbol{v}\rangle=0$.
2. 依次利用公理1、2, 再利用公理1, 可得: $\langle \boldsymbol{u},\boldsymbol{v}+\boldsymbol{w}\rangle=\langle \boldsymbol{v}+\boldsymbol{w},\boldsymbol{u}\rangle=\langle \boldsymbol{v},\boldsymbol{u}\rangle+\langle \boldsymbol{w},\boldsymbol{u}\rangle=\langle \boldsymbol{u},\boldsymbol{v}\rangle+\langle \boldsymbol{u},\boldsymbol{w}\rangle$.

6.8 内积空间的应用

本节的例题介绍了6.7节定义的内积空间是怎样在实际应用中出现的. 第一个例题与本章实例介绍中北美大地基准的更新有关.

6.8.1 加权最小二乘法

设 $\boldsymbol{y}$ 是含 n 个观测量 $y_1,\cdots,y_n$ 的向量. 假定我们希望用 $\mathbf{R}^n$ 的某指定子空间中的向量 $\hat{\boldsymbol{y}}$ 来逼近 $\boldsymbol{y}$(在6.5节中, $\hat{\boldsymbol{y}}$ 被写作 $A\boldsymbol{x}$, 从而 $\hat{\boldsymbol{y}}$ 属于 A 的列空间). 记 $\hat{\boldsymbol{y}}$ 中的元素为 $\hat{y}_1,\cdots,\hat{y}_n$. 则用 $\hat{\boldsymbol{y}}$ 逼近 $\boldsymbol{y}$ 时, 误差平方和或SS(E)是 [436]

$$\mathrm{SS(E)}=(y_1-\hat{y}_1)^2+\cdots+(y_n-\hat{y}_n)^2 \tag{1}$$

利用 $\mathbf{R}^n$ 中标准长度的记法, 上式可简写为 $\|\boldsymbol{y}-\hat{\boldsymbol{y}}\|^2$.

现在假设对 $\boldsymbol{y}$ 中各元素的观测并非同等可靠(对北美大地基准来说, 由于测量的时间跨度长达140年, 情况正是如此. 另外举个例子, $\boldsymbol{y}$ 中元素是用各式各样大小不等的测量样本计算出来的). 于是我们需要权衡(1)中的各项误差平方, 使得: 测量数据越可靠, 所赋予的权值越大.[1]若以 w_1^2, $\cdots$, w_n^2 记权值, 则加权误差平方和是

$$\text{加权的 SS(E)}=w_1^2(y_1-\hat{y}_1)^2+\cdots+w_n^2(y_n-\hat{y}_n)^2 \tag{2}$$

上式就是 $\boldsymbol{y}-\hat{\boldsymbol{y}}$ 长度的平方, 其中长度由类似于6.7节例题1的内积导出, 即

$$\langle \boldsymbol{x},\boldsymbol{y}\rangle=w_1^2x_1y_1+\cdots+w_n^2x_ny_n$$

加权最小二乘问题有时可以方便的转化为等价的普通最小二乘问题. 设 $\boldsymbol{W}$ 是对角线元素为(正数) $w_1,\cdots,w_n$ 的对角阵, 使得

$$\boldsymbol{W}\boldsymbol{y}=\begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & & \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & w_n \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}=\begin{bmatrix} w_1y_1 \\ w_2y_2 \\ \vdots \\ w_ny_n \end{bmatrix}$$

---

1. 有统计学知识的读者请注意: 假设测量的误差是独立的随机变量, 且均值为零, 方差为 $\sigma_1^2,\cdots,\sigma_n^2$. 则(2)中适当的权值是 $w_i^2=1/\sigma_i^2$. 误差的方差越大, 权值就越小.

与 $W\hat{y}$ 的表达式类似. 注意(2)的第 j 项可以写成

$$w_j^2(y_j-\hat{y}_j)^2=(w_jy_j-w_j\hat{y}_j)^2$$

这就表明(2)中的加权 SS(E) 等于 $\mathbf{R}^n$ 中 $W\boldsymbol{y}-W\hat{\boldsymbol{y}}$ 普通长度的平方, 可记为 $\|W\boldsymbol{y}-W\hat{\boldsymbol{y}}\|^2$.

现在假设逼近向量要由矩阵 A 的列来构造. 那么我们要求 $\hat{\boldsymbol{x}}$, 使得 $A\hat{\boldsymbol{x}}=\hat{\boldsymbol{y}}$ 尽可能地接近 $\boldsymbol{y}$. 然而, 接近程度由下面的加权误差来度量.

$$\|W\boldsymbol{y}-W\hat{\boldsymbol{y}}\|^2=\|W\boldsymbol{y}-WA\hat{\boldsymbol{x}}\|^2$$

这样, $\hat{\boldsymbol{x}}$ 是方程

437

$$WA\boldsymbol{x}=W\boldsymbol{y}$$

的(普通)最小二乘解. 其对应的正规方程是:

$$(WA)^TWA\boldsymbol{x}=(WA)^TW\boldsymbol{y}$$

【例题 1】 求最佳拟合数据(−2,3), (−1,5), (0,5), (1,4)和(2,3)的最小二乘直线 $y=\beta_0+\beta_1x$. 假设后两个数据点中 y 值的测量误差大于其他点, 将这两个数据的权值设置成其余数据的一半.

解: 与6.5节中一样, 以 X 记矩阵 A, 以 $\boldsymbol{\beta}$ 记向量 $\boldsymbol{x}$, 我们有:

$$X=\begin{bmatrix}1&-2\\1&-1\\1&0\\1&1\\1&2\end{bmatrix},\quad \boldsymbol{\beta}=\begin{bmatrix}\beta_0\\\beta_1\end{bmatrix},\quad \boldsymbol{y}=\begin{bmatrix}3\\5\\5\\4\\3\end{bmatrix}$$

对于权值矩阵 W, 令其对角线元素为2, 2, 2, 1和1. 通过左乘 W, 缩放 X 和 $\boldsymbol{y}$ 的行:

$$WX=\begin{bmatrix}2&-4\\2&-2\\2&0\\1&1\\1&2\end{bmatrix},\quad W\boldsymbol{y}=\begin{bmatrix}6\\10\\10\\4\\3\end{bmatrix}$$

对于正规方程, 计算

$$(WX)^TWX=\begin{bmatrix}14&-9\\-9&25\end{bmatrix}\quad 和\quad (WX)^TW\boldsymbol{y}=\begin{bmatrix}59\\-34\end{bmatrix}$$

并求解

$$\begin{bmatrix}14&-9\\-9&25\end{bmatrix}\begin{bmatrix}\beta_0\\\beta_1\end{bmatrix}=\begin{bmatrix}59\\-34\end{bmatrix}$$

正规方程的解（保留两位有效数字）是 $\beta_0=4.3$, $\beta_1=0.20$. 得到的直线是:

$$y=4.3+0.20x$$

相反, 这些数据的普通最小二乘直线是

$$y=4.0-0.10x$$

图6-39描绘了这两条直线. ■

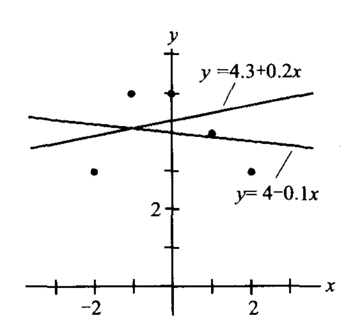


图6-39 加权最小二乘直线与普通最小二乘直线

6.8.2 数据趋势分析

设f是一个未知函数，它在$t_0,\cdots,t_n$处的值已知(有可能只是近似值). 如果数据$f(t_0),\cdots,f(t_n)$呈现“线性趋势”，那么我们可以期望用形如$\beta_0+\beta_1 t$的函数逼近f的值. 如果数据呈现“二次趋势”，那么我们应该尝试用形如$\beta_0+\beta_1 t+\beta_2 t^2$的函数. 对此，我们早在6.6节中已从另一角度进行了讨论.

438

在某些统计学问题中，将线性趋势与二次趋势(可能还有三次或者更高的趋势)区分开十分重要. 例如，假定工程师正在分析一辆新汽车的性能，$f(t)$代表t时刻汽车与某参考点间的距离. 如果汽车匀速行驶，则$f(t)$的图像应该是一条直线，其斜率为汽车行驶速度. 如果突然将油门踩到底，$f(t)$的图像将发生变化，它将包含一个二次项也许还有一个三次项(由加速产生的). 又比如，为了分析这辆汽车超越另一辆车的能力，工程师需要将这些二次项、三次项从线性项中分离出来.

如果用形如$y=\beta_0+\beta_1 t+\beta_2 t^2$的曲线来逼近未知函数，由于在统计意义上系数$\beta_2$不一定“独立”于其他$\beta_i$，所以它不一定能给出我们所期望的数据中二次趋势的有关信息. 为了进行所谓的数据**趋势分析**(trend analysis)，我们在$\mathbf{P}_n$上引入类似于6.7节例题2的内积. 对p，$q\in\mathbf{P}_n$，定义：

$$\langle p,q\rangle=p(t_0)q(t_0)+\cdots+p(t_n)q(t_n)$$

实际应用中，统计学家们很少需要考虑数据中高于三次或四次的趋势. 因此，用p_0，p_1，p_2和p_3记$\mathbf{P}_n$子空间$\mathbf{P}_3$的一组正交基，它通过对多项式1，t，t^2和t^3运用格拉姆-施密特方法而得. 由第2章补充题11可知，$\mathbf{P}_n$中存在多项式g，g在$t_0,\cdots,t_n$处的值与未知函数一致. 令$\hat{g}$为g在$\mathbf{P}_3$上的正交投影(相对于给定的内积)，比方说，

$$\hat{g}=c_0p_0+c_1p_1+c_2p_2+c_3p_3$$

则称$\hat{g}$为三次**趋势函数**(trend function)，称c_0，…，c_3为数据的**趋势系数**(trend coefficients). 趋势系数c_1，c_2和c_3分别度量线性趋势、二次趋势和三次趋势. 结果表明，如果数据具有某种性质，这些趋势系数在统计意义上相互独立.

由于p_0，p_1，p_2和p_3相互正交，趋势系数可以被逐个计算出来，并且是相互独立的(回忆$c_i=\langle g,p_i\rangle/\langle p_i,p_i\rangle$). 如果只需要二次趋势，我们就可以忽略$p_3$和$c_3$. 另外举个例子，如果需要确定四次趋势，我们就只要（用格拉姆-施密特方法）找$\mathbf{P}_4$中一个正交于$\mathbf{P}_3$的多项式p_4，并计算$\langle g,p_4\rangle/\langle p_4,p_4\rangle$.

【例题2】 t_0，…，t_n呈等间隔分布并且总和为零，是趋势分析最简单而常用的情

形. 试用一个二次趋势函数拟合数据$(-2,3)$, $(-1,5)$, $(0,5)$, $(1,4)$和$(2,3)$.

439

解: 这些 t-坐标的分布刚好使我们能利用6.7节例题5的结论, 得到

$$\begin{array}{lcccc} \text{多项式:} & p_0 & p_1 & p_2 & \text{数据}: g \\ \text{值向量:} & \begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}, & \begin{bmatrix}-2\\-1\\0\\1\\2\end{bmatrix}, & \begin{bmatrix}2\\-1\\-2\\-1\\2\end{bmatrix}, & \begin{bmatrix}3\\5\\5\\4\\3\end{bmatrix} \end{array}$$

计算过程只涉及这些向量, 与正交多项式的特定公式无关: $\mathbf{P}_2$中多项式对数据的最佳逼近是下面给出的正交投影

$$\hat{p}=\frac{\langle g,p_0\rangle}{\langle p_0,p_0\rangle}p_0+\frac{\langle g,p_1\rangle}{\langle p_1,p_1\rangle}p_1+\frac{\langle g,p_2\rangle}{\langle p_2,p_2\rangle}p_2=\frac{20}{5}p_0-\frac{1}{10}p_1-\frac{7}{14}p_2$$

及

$$\hat{p}(t)=4-0.1t-0.5(t^2-2) \tag{3}$$

由于p_2的系数并不是非常小, 我们可以合理地得出结论: 数据至少呈现二次趋势. 图6-40中所绘的图像也确认了这一点. ■

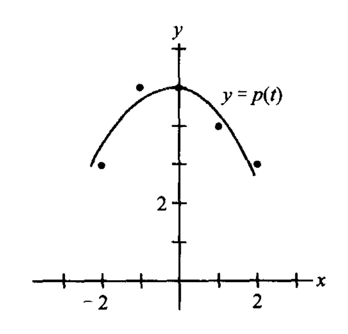


图6-40　使用二次趋势函数进行逼近

6.8.3　傅里叶级数(需要微积分知识)

连续函数通常可以用正弦和余弦函数的线性组合来逼近. 例如, 表示声波、某种电子信号或机械系统振动的函数.

简单起见, 我们考虑$0\leqslant t\leqslant 2\pi$上的函数. 事实是$C[0,2\pi]$上的任意函数都可以用如下形式的函数充分逼近:

$$\frac{a_0}{2}+a_1\cos t+\cdots+a_n\cos nt+b_1\sin t+\cdots+b_n\sin nt \tag{4}$$

其中n取值充分大. 函数(4)称作**三角多项式**(trigonometric polynomial). 如果a_n和b_n不同时为零, 我们说这个多项式是**n阶**(order n)的. 三角多项式与其他函数之间的联系依赖于下述事实: 对任意$n\geqslant 1$, 集合

$$\{1,\cos t,\cos 2t,\cdots,\cos nt,\sin t\ \sin 2t,\cdots,\sin nt\} \tag{5}$$

相对于内积

$$\langle f,g\rangle = \int_0^{2\pi} f(t)g(t)\,dt \tag{6}$$

是正交的. 其正交性将在下面的例题以及习题 5 和 6 中得到验证.

【例题 3】 设 $C[0,2\pi]$ 上有内积(6), 且令 m, n 是不相等的正整数. 证明 $\cos mt$ 与 $\sin nt$ 正交.

440

解: 利用三角恒等式. 当 $m\neq n$ 时, 有:

$$\begin{aligned}\langle \cos mt, \cos nt\rangle &= \int_0^{2\pi} \cos mt \cos nt\,dt\\ &= \frac{1}{2}\int_0^{2\pi} [\cos(mt+nt)+\cos(mt-nt)]\,dt\\ &= \frac{1}{2}\left[\frac{\sin(mt+nt)}{m+n}+\frac{\sin(mt-nt)}{m-n}\right]\Bigg|_0^{2\pi}=0\end{aligned}$$

■

设 W 是由(5)中函数所张成的 $C[0,2\pi]$ 的子空间. 给定 $f\in C[0,2\pi]$, W 中函数对 f 的最佳逼近称为 $[0,2\pi]$ 上 f 的 **n 阶傅里叶逼近**(*n*th-order Fourier approximation). 因为(5)中的函数是正交的, 这个最佳逼近就是 W 上的正交投影. 此时, (4)中的系数 a_k 和 b_k 称为 f 的**傅里叶系数**(Fourier coefficient). 正交投影的标准公式表明

$$a_k=\frac{\langle f,\ \cos kt\rangle}{\langle \cos kt,\ \cos kt\rangle},\quad b_k=\frac{\langle f,\ \sin kt\rangle}{\langle \sin kt,\ \sin kt\rangle},\quad k\geqslant 1$$

习题 7 要求证明 $\langle \cos kt, \cos kt\rangle=\pi$ 和 $\langle \sin kt, \sin kt\rangle=\pi$. 于是

$$a_k=\frac{1}{\pi}\int_0^{2\pi} f(t)\cos kt\,dt,\quad b_k=\frac{1}{\pi}\int_0^{2\pi} f(t)\sin kt\,dt \tag{7}$$

正交投影中(常)函数 1 的系数为:

$$\frac{\langle f,1\rangle}{\langle 1,1\rangle}=\frac{1}{2\pi}\int_0^{2\pi} f(t)\cdot 1\,dt=\frac{1}{2}\left[\frac{1}{\pi}\int_0^{2\pi} f(t)\cos(0\cdot t)\,dt\right]=\frac{a_0}{2}$$

其中对于 $k=0$, a_0 由(7)定义. 这就解释了(4)中常数项为什么是 $a_0/2$.

【例题 4】 求函数 $f(t)=t$ 在区间 $[0,2\pi]$ 上的 n 阶傅里叶逼近.

解: 我们计算

$$\frac{a_0}{2}=\frac{1}{2}\cdot\frac{1}{\pi}\int_0^{2\pi} t\,dt=\frac{1}{2\pi}\left[\frac{1}{2}t^2\Big|_0^{2\pi}\right]=\pi$$

并且对 $k>0$, 利用分部积分法

$$a_k=\frac{1}{\pi}\int_0^{2\pi} t\cos kt\,dt=\frac{1}{\pi}\left[\frac{1}{k^2}\cos kt+\frac{t}{k}\sin kt\right]_0^{2\pi}=0$$

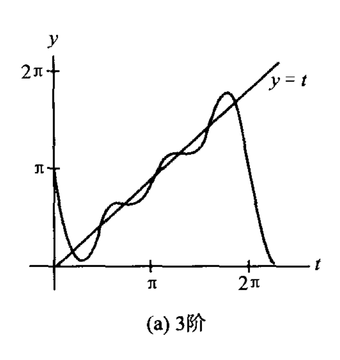
441

$$b_k=\frac{1}{\pi}\int_0^{2\pi} t\sin kt\,dt=\frac{1}{\pi}\left[\frac{1}{k^2}\sin kt-\frac{t}{k}\cos kt\right]_0^{2\pi}=-\frac{2}{k}$$

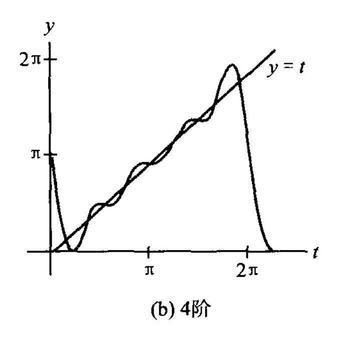
这样, $f(t)=t$ 的 n 阶傅里叶逼近是

$$\pi-2\sin t-\sin 2t-\frac{2}{3}\sin 3t-\cdots-\frac{2}{n}\sin nt$$

图 6-41 给出了 f 的 3 阶和 4 阶傅里叶逼近. ■



(a) 3阶



(b) 4阶

图6-41　函数$f(t)=t$的傅里叶逼近

f与一个傅里叶逼近间距离的范式称作**均方误差**(mean square error)(用"均"这个字是因为这个范式由积分确定). 可以证明，当傅里叶逼近的阶数增加时，均方误差趋近于零. 因此，我们通常记

$$f(t)=\frac{a_0}{2}+\sum_{m=1}^{\infty}(a_m\cos mt+b_m\sin mt)$$

$f(t)$的这个表达式称为f在$[0,2\pi]$上的**傅里叶级数**(Fourier series). 例如，项$a_m\cos mt$就是f在由$\cos mt$张成的1维子空间上的投影.

基础练习

1. 设$q_1(t)=1$，$q_2(t)=t$及$q_3(t)=3t^2-4$. 验证$\{q_1,q_2,q_3\}$是$C[-2,2]$的正交集，空间上的内积由6.7节例题7给出(从-2到2进行积分).
2. 求下列函数的1阶和3阶傅里叶逼近：

442

$$f(t)=3-2\sin t+5\sin 2t-6\cos 2t$$

习题　6.8

1. 求最佳拟合数据(-2,0)，(-1,0)，(0,2)，(1,4)和(2,4)的最小二乘直线$y=\beta_0+\beta_1x$. 假定第一个以及最后一个数据点不可靠，将它们的权值设置成其他点的一半.
2. 假设在一个加权最小二乘问题中，25个数据点中有5个点的y测量值不及其他点可靠，要将它们的权值设置成其余20个点的一半. 一种方法是将那20个点的权值乘以因子1，而余下5个点则乘以因子$\frac{1}{2}$. 第二种方法是将那20个点的权值乘以因子2，而余下5个点则乘以因子1. 这两种方法会不会得到不同的结果？请予以解释.
3. 用三次趋势函数拟合例题2中的数据. 正交三次多项式为$p_3(t)=\frac{5}{6}t^3-\frac{17}{6}t$.
4. 为了对6个等间隔分布的数据点进行趋势分析，我们可以利用与$t=-5$，-3，-1，1，3和5各点赋值相关的正交多项式.

   a. 证明前三个正交多项式是：

$$p_0(t)=1,p_1(t)=t,\text{and }p_2(t)=\frac{3}{8}t^2-\frac{35}{8}$$

   (正交多项式被适当缩放,以使其在赋值点处的取值为小整数).

   b. 用二次趋势函数拟合数据

$(-5,1)$, $(-3,1)$, $(-1,4)$, $(1,4)$, $(3,6)$, $(5,8)$

习题5~14中，空间 $C[0,2\pi]$具有内积(6).

5. 证明：当 $m \neq n$ 时，$\sin mt$ 与 $\sin nt$ 正交.
6. 证明：对任意正整数 m，n，$\sin mt$ 与 $\cos nt$ 正交.
7. 证明：对 $k>0$，$\|\cos kt\|^2 = \pi$且$\|\sin kt\|^2 = \pi$.
8. 求$f(t) = t-1$ 的 3 阶傅里叶逼近.
9. 求$f(t) = 2\pi - t$ 的 3 阶傅里叶逼近.
10. 求方波函数$f(t)$的 3 阶傅里叶逼近. 其中$0 \leqslant t < \pi$时$f(t) = 1$，$\pi \leqslant t < 2\pi$时$f(t) = -1$.
11. 不用积分运算，求 $\sin^2 t$ 的 3 阶傅里叶逼近.
12. 不用积分运算，求 $\cos^3 t$ 的 3 阶傅里叶逼近.
13. 解释为什么两个函数之和的傅里叶系数等于两函数对应的傅里叶系数之和.
14. 假设 $C[0,2\pi]$中某函数f的前几个傅里叶系数为 a_0，a_1，a_2和 b_1，b_2，b_3. 下列哪个三角多项式更接近于f？说明理由.

$$g(t) = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + b_1 \sin t$$

$$h(t) = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + b_1 \sin t + b_2 \sin 2t$$

15. [**M**]参考6.6节习题13中关于飞机起飞性能的数据. 假设随着飞机速度的增加，测量误差增大，并且令 W 是对角线元素为1，1，1，0.9，0.9，0.8，0.7，0.6，0.5，0.4，0.3，0.2，0.1的对角权值矩阵. 求三次曲线，使之用最小的加权最小二乘误差拟合数据. 并且用它估计飞机在 $t=4.5$ 秒时的速度.
16. [**M**]设f_4和f_5分别是习题10中方波函数在 $C[0,2\pi]$中的4阶和5阶逼近. 分别绘出f_4和f_5在区间$[0,2\pi]$上，以及f_5在$[-2\pi,2\pi]$上的图像.

基础练习答案

1. 计算

$$\langle q_1, q_2 \rangle = \int_{-2}^{2} 1 \cdot t\,dt = \frac{1}{2}t^2 \Big|_{-2}^{2} = 0$$

$$\langle q_1, q_3 \rangle = \int_{-2}^{2} 1 \cdot (3t^2 - 4)\,dt = (t^3 - 4t) \Big|_{-2}^{2} = 0$$

$$\langle q_2, q_3 \rangle = \int_{-2}^{2} t \cdot (3t^2 - 4)\,dt = \left(\frac{3}{4}t^4 - 2t^2\right) \Big|_{-2}^{2} = 0$$

443

2. f的 3 阶傅里叶逼近是 $C[0,2\pi]$空间中用 1，$\cos t$，$\cos 2t$，$\cos 3t$，$\sin t$，$\sin 2t$ 和 $\sin 3t$ 所张成子空间内函数(向量)对f的最佳逼近(见图6-42). 但是，f显然属于这个子空间，所以f就是自己的最佳逼近：

$$f(t) = 3 - 2\sin t + 5\sin 2t - 6\cos 2t$$

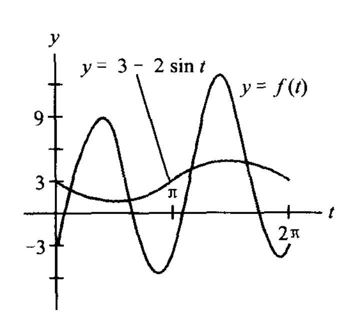


图6-42 $f(t)$的1阶和3阶逼近

对于1阶逼近，子空间 $W = \mathrm{Span}\{1, \cos t, \sin t\}$中与$f$距离最近的函数是 $3 - 2\sin t$. $f(t)$公式中的其余两项与 W 中函数正交，所以对于1阶逼近，这两项不会影响傅里叶系数的积分结果.

第 6 章补充题

1. 下列命题所讨论的都是具有标准内积的 $\mathbf{R}^n$(或 $\mathbf{R}^m$)空间中的向量. 判断各命题的真假，并

说明理由.

a. 向量的长度是一个正数.

b. 向量 $\boldsymbol{v}$ 与负向量 $-\boldsymbol{v}$ 长度相等.

c. $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 之间的距离是 $\|\boldsymbol{u}-\boldsymbol{v}\|$.

d. 如果 r 是任意一个数量，则 $\|r\boldsymbol{v}\|=r\|\boldsymbol{v}\|$.

e. 如果两向量正交，则它们线性无关.

f. 如果 $\boldsymbol{x}$ 同时与 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 正交，则 $\boldsymbol{x}$ 一定正交于 $\boldsymbol{u}-\boldsymbol{v}$.

g. 若 $\|\boldsymbol{u}+\boldsymbol{v}\|^2=\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2$，则 $\boldsymbol{u}$ 与 $\boldsymbol{v}$ 正交.

h. 若 $\|\boldsymbol{u}-\boldsymbol{v}\|^2=\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2$，则 $\boldsymbol{u}$ 与 $\boldsymbol{v}$ 正交.

i. $\boldsymbol{y}$ 在 $\boldsymbol{u}$ 上的正交投影是 $\boldsymbol{y}$ 的数量倍.

j. 如果向量 $\boldsymbol{y}$ 在子空间 W 上的正交投影就是它本身，则 $\boldsymbol{y}$ 属于 W.

k. $\mathbf{R}^n$ 中与某指定向量正交的全体向量作成的集合是 $\mathbf{R}^n$ 的一个子空间.

l. 如果 W 是 $\mathbf{R}^n$ 的子空间，则 W 与 $W^\perp$ 没有公共向量.

m. 如果 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 是正交集，c_1，c_2，c_3 为数量，则 $\{c_1\boldsymbol{v}_1,c_2\boldsymbol{v}_2,c_3\boldsymbol{v}_3\}$ 仍是正交集.

n. 如果矩阵 U 有标准正交列，则 $UU^{\mathrm{T}}=I$.

o. 各列正交的方阵是正交矩阵.

p. 如果某方阵有标准正交列，则其各行亦标准正交.

q. 若 W 是子空间，则 $\|\mathrm{proj}_W\boldsymbol{v}\|^2+\|\boldsymbol{v}-\mathrm{proj}_W\boldsymbol{v}\|^2=\|\boldsymbol{v}\|^2$.

r. $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解是 Col A 中与 $\boldsymbol{b}$ 距离最近的向量 $A\hat{\boldsymbol{x}}$，使得对任意 $\boldsymbol{x}$，$\|\boldsymbol{b}-A\hat{\boldsymbol{x}}\|\leqslant\|\boldsymbol{b}-A\boldsymbol{x}\|$.

s. $A\boldsymbol{x}=\boldsymbol{b}$ 最小二乘解的正规方程由 $\hat{\boldsymbol{x}}=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b}$ 给出.

2. 设 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是一个标准正交集. 从 $p=2$ 开始归纳验证下面的等式. 若 $\boldsymbol{x}=c_1\boldsymbol{v}_1+\cdots+c_p\boldsymbol{v}_p$，则

$$\|\boldsymbol{x}\|^2=|c_1|^2+\cdots+|c_p|^2$$

3. 设 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 是 $\mathbf{R}^n$ 的一个标准正交集. 验证下列被称为贝塞尔不等式的不等式，它对 $\mathbf{R}^n$ 中任意 $\boldsymbol{x}$ 都成立：

$$\|\boldsymbol{x}\|^2\geqslant|\boldsymbol{x}\cdot\boldsymbol{v}_1|^2+|\boldsymbol{x}\cdot\boldsymbol{v}_2|^2+\cdots+|\boldsymbol{x}\cdot\boldsymbol{v}_p|^2$$

4. 设 U 是一个 $n\times n$ 正交矩阵. 证明：若 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$ 是的一组标准正交基，则 $\{U\boldsymbol{v}_1,\cdots,U\boldsymbol{v}_n\}$ 亦然.

5. 证明：如果对 $\mathbf{R}^n$ 中的任意 $\boldsymbol{x}$ 和 $\boldsymbol{y}$，矩阵 U 满足 $(U\boldsymbol{x})\cdot(U\boldsymbol{y})=\boldsymbol{x}\cdot\boldsymbol{y}$，则 U 是正交矩阵.

6. 证明：如果 U 是一个正交矩阵，则 U 的实本征值一定是 ±1.

7. 一个豪斯霍尔德矩阵或初等反射形如 $Q=I-2\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}$，其中 $\boldsymbol{u}$ 为单位向量（见第2章补充题第13题）. 证明 Q 是正交矩阵.（计算机程序中常用初等反射生成矩阵 A 的 QR 分解. 若 A 的各列线性无关，则将 A 左乘一系列初等反射后可得到一个上三角矩阵.）

444

8. 设 $T:\mathbf{R}^n\to\mathbf{R}^n$ 是保持长度的线性变换，即对 $\mathbf{R}^n$ 中任意 $\boldsymbol{x}$，有 $\|T(\boldsymbol{x})\|=\|\boldsymbol{x}\|$.

   a. 证明 T 也保持正交性，即 $\boldsymbol{x}\cdot\boldsymbol{y}=0$ 时有 $T(\boldsymbol{x})\cdot T(\boldsymbol{y})=0$.

   b. 证明 T 的标准矩阵是正交矩阵.

9. 设 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是 $\mathbf{R}^n$ 中线性无关但不正交的向量. 试说明：若不事先构造 Span$\{\boldsymbol{u},\boldsymbol{v}\}$，怎样求 $x_1\boldsymbol{u}+x_2\boldsymbol{v}$ 形式的向量对 $\mathbf{R}^n$ 中向量 $\boldsymbol{z}$ 的最佳逼近?

10. 假定 A 的各列线性无关. 试说明：用 $c\boldsymbol{b}$（c 为某一非零数量）代替 $\boldsymbol{b}$ 时，$A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解 $\hat{\boldsymbol{x}}$ 将有何变化?

11. 若 a，b 和 c 互不相等，则下列方程组中方程的图像互相平行，因此方程组不相容. 证明：这个方程组最小二乘解的集合恰好是方程 $x-2y+5z=(a+b+c)/3$ 所表示的平面.

$x-2y+5z=a$
$x-2y+5z=b$
$x-2y+5z=c$

12. 若已知 $n\times n$ 矩阵 $\boldsymbol{A}$ 的一个近似本征向量 $\boldsymbol{v}$，考虑如何求 $\boldsymbol{A}$ 的一个本征值. 由于 $\boldsymbol{v}$ 并不精确，因此方程

$$A\boldsymbol{v}=\lambda\boldsymbol{v} \tag{1}$$

可能没有解. 不过，适当观察(1)，我们就能发现 λ 可以用最小二乘解估计出来. 把 $\boldsymbol{v}$ 看成一个 $n\times 1$ 矩阵 V，把 $\boldsymbol{\lambda}$ 看成 $\mathbf{R}^1$ 中的向量，并且用符号 $\boldsymbol{b}$ 记向量 $A\boldsymbol{v}$. 这样（1）就变成了 $\boldsymbol{b}=\boldsymbol{\lambda}V$，也可写作 $V\boldsymbol{\lambda}=\boldsymbol{b}$. 求解这个含 n 个方程、一个未知量 $\boldsymbol{\lambda}$ 的方程组的最小二乘解，并将它用原始记号写出来. $\boldsymbol{\lambda}$ 的这个估计结果称作瑞利商. 见 5.8 节习题 11 和 12.

13. 按照下面步骤，证明 $m\times n$ 矩阵 $\boldsymbol{A}$ 所确定的四个基本空间之间的如下关系：

$$\text{Row } A=(\text{Nul } A)^{\perp}, \text{Col } A=(\text{Nul } A^{\mathrm{T}})^{\perp}$$

a. 证明 Row $\boldsymbol{A}$ 包含于(Nul $\boldsymbol{A}$)$^{\perp}$.（证明如果 $\boldsymbol{x}\in$ Row $\boldsymbol{A}$，则 $\boldsymbol{x}$ 与 Nul $\boldsymbol{A}$ 中的任意 $\boldsymbol{u}$ 正交.）

b. 设 rank $\boldsymbol{A}=r$. 求 dim Nul $\boldsymbol{A}$ 和 dim（Nul $\boldsymbol{A}$）$^{\perp}$，并且根据(a)推导 Row $\boldsymbol{A}$ =（Nul $\boldsymbol{A}$）$^{\perp}$. [提示：研究 6.3 节中的习题.]

c. 说明 Col $\boldsymbol{A}$ =（Nul $\boldsymbol{A}^{\mathrm{T}}$）$^{\perp}$ 的理由.

14. 解释为什么方程 $A\boldsymbol{x}=\boldsymbol{b}$ 有解当且仅当 $\boldsymbol{b}$ 与方程 $A^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{0}$ 的所有解都正交.

习题 15 和 16 涉及 $n\times n$ 矩阵 $\boldsymbol{A}$ 的(实)舒尔分解 $A=URU^{\mathrm{T}}$，其中 $\boldsymbol{U}$ 是正交矩阵，$\boldsymbol{R}$ 是 $n\times n$ 上三角矩阵.[1]

15. 证明：如果 $\boldsymbol{A}$ 有(实)舒尔分解 $A=URU^{\mathrm{T}}$，则计重数，$\boldsymbol{A}$ 有 n 个实本征值.

16. 设 $\boldsymbol{A}$ 是有 n 个（计重数）实本征值的 $n\times n$ 矩阵，记作 $\lambda_1,\cdots,\lambda_n$. 可以证明 $\boldsymbol{A}$ 有(实)舒尔分解. (a)和(b)提供了证明的关键思想. 为完成整个证明，只需对小矩阵连续运用(a)、(b)，最后综合起来即可.

(a) 设 $\boldsymbol{u}_1$ 是对应于 λ_1 的单位本征向量，任取 $\boldsymbol{u}_1$ 以外的其他向量 $\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$，使 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n\}$ 构成 $\mathbf{R}^n$ 的一组正交基，令 $U=[\boldsymbol{u}_1\ \ \boldsymbol{u}_2\ \ \cdots\ \ \boldsymbol{u}_n]$. 证明：$U^{\mathrm{T}}AU$ 的首列是 $\lambda_1\boldsymbol{e}_1$，其中 $\boldsymbol{e}_1$ 是单位矩阵的首列.

(b) (a)表明具有下面的形式. 解释为什么 $\boldsymbol{A}_1$ 的本征值是 $\lambda_2,\cdots,\lambda_n$. [提示：参考第 5 章补充题.]

$$U^{\mathrm{T}}AU=\begin{bmatrix}\lambda_1 & * & * & * & * \\ 0 & & & & \\ \vdots & & A_1 & & \\ 0 & & & & \end{bmatrix}$$

[M]稍微改变方程 $A\boldsymbol{x}=\boldsymbol{b}$ 的右端，比如说，变成 $A\boldsymbol{x}=\boldsymbol{b}+\Delta\boldsymbol{b}$，其中 $\Delta\boldsymbol{b}$ 是某一向量，那么方程的解将从 $\boldsymbol{x}$ 变成 $\boldsymbol{x}+\Delta\boldsymbol{x}$，其中 $\Delta\boldsymbol{x}$ 满足 $A\Delta\boldsymbol{x}=\Delta\boldsymbol{b}$. 商 $\|\Delta\boldsymbol{b}\|/\|\boldsymbol{b}\|$ 称作 $\boldsymbol{b}$ 的**相对变化**(relative change)（当 $\Delta\boldsymbol{b}$ 表示 $\boldsymbol{b}$ 中元素的误差时，这个商也称作 $\boldsymbol{b}$ 的**相对误差**(relative error)）. 解的相对变化是 $\|\Delta\boldsymbol{x}\|/\|\boldsymbol{x}\|$. $\boldsymbol{A}$ 可逆时，$\boldsymbol{A}$ 的**条件数**(condition number)，记作 cond($\boldsymbol{A}$)，为 x 的相对变化提供了一个上界：

445

$$\frac{\|\Delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|}\leqslant \text{cond}(A)\cdot\frac{\|\Delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|} \tag{2}$$

1. 若允许有复数，每个 $n\times n$ 矩阵 $\boldsymbol{A}$ 都有(复)舒尔分解 $A=URU^{-1}$，其中 $\boldsymbol{R}$ 是上三角矩阵，U^{-1} 是 $\boldsymbol{U}$ 的共轭转置. Roger A. Horn, Charles R. Johnson, *Matrix Analysis* (Cambridge: Cambridge University Press, 1985), pp. 79－100 中对这一有用的事实进行了讨论.

在习题17～20中，求解 $\boldsymbol{Ax}=\boldsymbol{b}$ 和 $\boldsymbol{A}(\Delta\boldsymbol{x})=\Delta\boldsymbol{b}$，并证明(2)式成立.（参考2.3节习题41～43中关于病态矩阵的讨论.）

17. $\boldsymbol{A}=\begin{bmatrix}4.5 & 3.1\\1.6 & 1.1\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}19.249\\6.843\end{bmatrix}$, $\Delta\boldsymbol{b}=\begin{bmatrix}0.001\\-0.003\end{bmatrix}$

18. $\boldsymbol{A}=\begin{bmatrix}4.5 & 3.1\\1.6 & 1.1\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}0.500\\-1.407\end{bmatrix}$, $\Delta\boldsymbol{b}=\begin{bmatrix}0.001\\-0.003\end{bmatrix}$

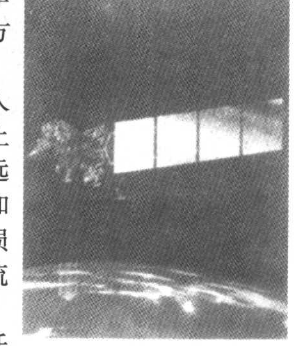
19. $\boldsymbol{A}=\begin{bmatrix}7 & -6 & -4 & 1\\-5 & 1 & 0 & -2\\10 & 11 & 7 & -3\\19 & 9 & 7 & 1\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}0.100\\2.888\\-1.404\\1.462\end{bmatrix}$, $\Delta\boldsymbol{b}=10^{-4}\begin{bmatrix}0.49\\-1.28\\5.78\\8.04\end{bmatrix}$

20. $\boldsymbol{A}=\begin{bmatrix}7 & -6 & -4 & 1\\-5 & 1 & 0 & -2\\10 & 11 & 7 & -3\\19 & 9 & 7 & 1\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}4.230\\-11.043\\49.991\\69.536\end{bmatrix}$, $\Delta\boldsymbol{b}=10^{-4}\begin{bmatrix}0.27\\7.76\\-3.77\\3.93\end{bmatrix}$

第7章 对称矩阵和二次型

实例介绍：多通道图像处理

大约每过80分钟，两颗地球资源探测卫星就会无声地划过天空，沿靠近极地的轨道环绕地球一周，并且以185公里的宽幅记录陆地和海岸的图像. 每颗卫星每16天会掠遍地球表面的几乎每一平方公里，所以地球上的每个角落每过8天就会被监测到一次.



探测卫星所记录的图像有多种用途. 城市开发和规划人员利用它们研究城市扩建、工业发展的速度和方向，以及土地利用的其他动态. 农村则可以用于分析土壤湿度，对偏远地区的植被分类，定位内陆的湖泊和溪流. 政府可以检测和评估自然灾害，如森林火灾、熔岩流、洪水和飓风造成的损失. 环保部门可以确定工业污染，测量工厂附近湖泊与河流的水温.

出于研究的需要，卫星上的传感器可同步获取地球上任意区域的七幅图像. 这些传感器记录不同波段的能量——其中三个属于可见光光谱，四个属于红外波段和热波段. 每幅图像经过数字化被存储为一个数值阵列，其中的数值代表图像中对应点(或*像素*)的信号强度. 七幅图中每一幅都是*多通道*或*多谱段图像*中的一个通道.

一个指定区域的七幅卫星图像通常包含许多冗余信息，这是因为某些图像特征可能同时出现在多幅图当中. 不过，还有些特征，因为颜色或者温度的缘故，所反射的光只能被一两个传感器所记录. 多通道图像处理的目的之一就是要压缩数据中的信息，而不是独立研究每幅图.

*主成分分析*是一种消除冗余信息的有效方法，它将原始数据中的大部分信息呈现在一两幅合成图像中. 粗略地说，其目标是求原始图像的一个特殊的线性组合，即这样一列权重，它们将每个像素在七幅图像中对应的值组合成一个新值. 这些权重的选取使得合成图像(称作*第一主成分*)中光强度的变化幅度，即影像方差，比每一幅原始图像中的都大. 依照7.5节将要介绍的方法，其他成分图像也可被构造出来.

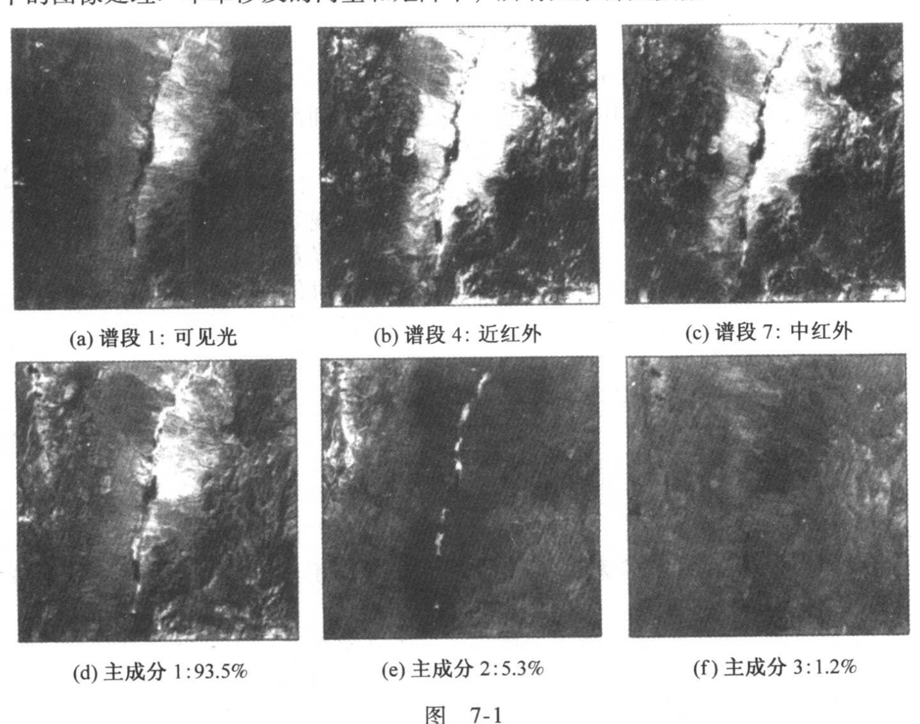
447

下面，我们以摄于内华达州铁路峡谷上空的照片为例，解释主成分分析. 探测卫星在三个谱段下得到的图像如图7-1a～c所示. 图7-1d～f中的三幅主成分图像将三个谱段的全部信息进行了重排. 第一个成分d显示(或者说“解释”)了原始数据中93.5%的影像方差. 使用这种方法，含三个通道的原始数据被压缩成单通道数据，仅损失了6.5%的影像方差.

位于马里兰州洛克维尔市的地球卫星公司友善地向我们提供上述照片，他们目前正对224个独立谱段下的图像进行试验. 对于这样庞大的数据集，主成分分析是十

448 分必要的，它通常可以将数据压缩成15个左右有用的主成分.

与其他各大类型的矩阵相比，对称矩阵在应用中更为普遍，并且呈现方式不一. 对称矩阵的理论丰富且极富美感，这在很大程度上依赖于第5章的对角化及第6章的正交性知识. 7.1节所介绍的对称矩阵的对角化是7.2节和7.3节二次型相关内容的基础. 而7.3节又为本章最后两节作了铺垫，那两节分别关于奇异值分解以及实例介绍中的图像处理. 本章涉及的向量和矩阵中，所有元素都是实数.



(a) 谱段1: 可见光　(b) 谱段4: 近红外　(c) 谱段7: 中红外

(d) 主成分1:93.5%　(e) 主成分2:5.3%　(f) 主成分3:1.2%

图　7-1

7.1　对称矩阵的对角化

对称(symmetric)矩阵是满足 $A^T=A$ 的一个矩阵 A. 这样的矩阵一定是方阵. 其主对角线元素是任意的，而其他元素则沿主对角线两侧对称出现.

【例题1】　下列各矩阵中，只有前三个是对称矩阵.

$$\text{对称：}\begin{bmatrix}1 & 0\\ 0 & -3\end{bmatrix},\begin{bmatrix}0 & -1 & 0\\ -1 & 5 & 8\\ 0 & 8 & -7\end{bmatrix},\begin{bmatrix}a & b & c\\ b & d & e\\ c & e & f\end{bmatrix}$$

$$\text{非对称：}\begin{bmatrix}1 & -3\\ 3 & 0\end{bmatrix},\begin{bmatrix}1 & -4 & 0\\ -6 & 1 & -4\\ 0 & -6 & 1\end{bmatrix},\begin{bmatrix}5 & 4 & 3 & 2\\ 4 & 3 & 2 & 1\\ 3 & 2 & 1 & 0\end{bmatrix}$$

■

在开始研究对称矩阵以前，回顾5.3节中的对角化方法是十分有益的.

【例题2】 如果可能，将矩阵 $\boldsymbol{A}=\begin{bmatrix}6&-2&-1\\-2&6&-1\\-1&-1&5\end{bmatrix}$ 对角化.

解： $\boldsymbol{A}$ 的特征方程是

$$0=-\lambda^3+17\lambda^2-90\lambda+144=-(\lambda-8)(\lambda-6)(\lambda-3)$$

通过标准的计算过程，可以得到每个本征空间的基：

$$\lambda=8:\ \boldsymbol{v}_1=\begin{bmatrix}-1\\1\\0\end{bmatrix};\qquad \lambda=6:\ \boldsymbol{v}_2=\begin{bmatrix}-1\\-1\\2\end{bmatrix};\qquad \lambda=3:\ \boldsymbol{v}_3=\begin{bmatrix}1\\1\\1\end{bmatrix}$$

这三个向量构成 $\mathbf{R}^3$ 的一组基，并且我们可以将它们作为矩阵 $\boldsymbol{P}$ 的列，使 $\boldsymbol{P}$ 对角化 $\boldsymbol{A}$. 不过，易见 $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3\}$ 是一个正交集，而且若 $\boldsymbol{P}$ 各列标准正交，则 $\boldsymbol{P}$ 会更有用. 因为本征向量的非零倍数仍然是本征向量，我们可以标准化 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 及 $\boldsymbol{v}_3$，以得到下列单位向量 449

$$\boldsymbol{u}_1=\begin{bmatrix}-1/\sqrt{2}\\1/\sqrt{2}\\0\end{bmatrix},\quad \boldsymbol{u}_2=\begin{bmatrix}-1/\sqrt{6}\\-1/\sqrt{6}\\2/\sqrt{6}\end{bmatrix},\quad \boldsymbol{u}_3=\begin{bmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$$

令

$$\boldsymbol{P}=\begin{bmatrix}-1/\sqrt{2}&-1/\sqrt{6}&1/\sqrt{3}\\1/\sqrt{2}&-1/\sqrt{6}&1\ \sqrt{3}\\0&2/\sqrt{6}&1/\sqrt{3}\end{bmatrix},\quad \boldsymbol{D}=\begin{bmatrix}8&0&0\\0&6&0\\0&0&3\end{bmatrix}$$

则与往常一样，有 $\boldsymbol{A}=\boldsymbol{PDP}^{-1}$. 不过，此处由于 $\boldsymbol{P}$ 是方阵且有标准正交列，所以 $\boldsymbol{P}$ 是正交矩阵，并且 $\boldsymbol{P}^{-1}$ 就是 $\boldsymbol{P}^{\mathrm{T}}$（见6.2节）. ■

定理1说明了例题2中的本征向量为什么正交——它们对应不同的本征值.

【定理1】 若 $\boldsymbol{A}$ 是对称矩阵，则任意两个取自不同本征空间的本征向量都正交.

证明： 设 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 是两个本征向量，它们对应不同的本征值，比如说 λ_1 和 λ_2. 为证明 $\boldsymbol{v}_1\cdot\boldsymbol{v}_2=0$，计算

$$\begin{aligned}\lambda_1\boldsymbol{v}_1\cdot\boldsymbol{v}_2&=(\lambda_1\boldsymbol{v}_1)^{\mathrm{T}}\boldsymbol{v}_2=(\boldsymbol{A}\boldsymbol{v}_1)^{\mathrm{T}}\boldsymbol{v}_2 &&\text{因为 }\boldsymbol{v}_1\text{ 是本征向量}\\&=(\boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}})\boldsymbol{v}_2=\boldsymbol{v}_1^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{v}_2) &&\text{因为 }\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}\\&=\boldsymbol{v}_1^{\mathrm{T}}(\lambda_2\boldsymbol{v}_2) &&\text{因为 }\boldsymbol{v}_2\text{ 是本征向量}\\&=\lambda_2\boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{v}_2=\lambda_2\boldsymbol{v}_1\cdot\boldsymbol{v}_2\end{aligned}$$

于是 $(\lambda_1-\lambda_2)\boldsymbol{v}_1\cdot\boldsymbol{v}_2=0$. 但 $\lambda_1-\lambda_2\neq0$，所以 $\boldsymbol{v}_1\cdot\boldsymbol{v}_2=0$. ■

例题2中特殊形式的对角化在对称矩阵理论中十分重要. 我们称矩阵 $\boldsymbol{A}$ **可正交对角化**（orthogonally diagonalizable），如果存在一个正交矩阵 $\boldsymbol{P}$（满足 $\boldsymbol{P}^{-1}=\boldsymbol{P}^{\mathrm{T}}$）和一个对角阵 $\boldsymbol{D}$ 使得

$$\boldsymbol{A}=\boldsymbol{PDP}^{\mathrm{T}}=\boldsymbol{PDP}^{-1}\tag{1}$$

为了正交对角化一个 $n\times n$ 矩阵，我们必须找出 n 个线性无关且标准正交的本征

向量. 这在何时是可能做到的？如果 $\boldsymbol{A}$ 是如(1)中所述的可正交对角化矩阵，则

$$\boldsymbol{A}^{\mathrm{T}}=(\boldsymbol{PDP}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{P}^{\mathrm{TT}}\boldsymbol{D}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}=\boldsymbol{PDP}^{\mathrm{T}}=\boldsymbol{A}$$

所以 $\boldsymbol{A}$ 是对称的！定理2表明：反过来，每个对称矩阵都可正交对角化. 定理的证明较难，在此省略. 定理3之后我们给出了一个证明的主要思想.

450

【定理2】　一个 $n\times n$ 矩阵 $\boldsymbol{A}$ 可正交对角化当且仅当 $\boldsymbol{A}$ 是对称矩阵.

这个定理十分出人意料. 因为第5章的经验告诉我们，通常要判别一个矩阵是否可对角化是不可能的. 不过，对于对称矩阵情况就不同了.

下面例题所处理的矩阵有重复的本征值.

【例题3】　正交对角化矩阵 $\boldsymbol{A}=\begin{bmatrix}3&-2&4\\-2&6&2\\4&2&3\end{bmatrix}$，其特征方程为

$$0=-\lambda^3+12\lambda^2-21\lambda-98=-(\lambda-7)^2(\lambda+2)$$

解： 经过标准的演算，得到本征空间的基：

$$\lambda=7:\quad \boldsymbol{v}_1=\begin{bmatrix}1\\0\\1\end{bmatrix},\boldsymbol{v}_2=\begin{bmatrix}-1/2\\1\\0\end{bmatrix};\qquad \lambda=-2:\boldsymbol{v}_3=\begin{bmatrix}-1\\-1/2\\1\end{bmatrix}$$

$\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 虽然线性无关，但不正交. 回忆6.2节，$\boldsymbol{v}_2$ 在 $\boldsymbol{v}_1$ 上的投影是 $\dfrac{\boldsymbol{v}_2\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1$，且 $\boldsymbol{v}_2$ 的正交于 $\boldsymbol{v}_1$ 的分量是

$$\boldsymbol{z}_2=\boldsymbol{v}_2-\frac{\boldsymbol{v}_2\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1=\begin{bmatrix}-1/2\\1\\0\end{bmatrix}-\frac{-1/2}{2}\begin{bmatrix}1\\0\\1\end{bmatrix}=\begin{bmatrix}-1/4\\1\\1/4\end{bmatrix}$$

所以 $\{\boldsymbol{v}_1,\boldsymbol{z}_2\}$ 是 $\lambda=7$ 所对应本征空间的一个正交集(注意 $\boldsymbol{z}_2$ 是本征向量 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的线性组合，所以 $\boldsymbol{z}_2$ 属于这个本征空间. $\boldsymbol{z}_2$ 的构造恰好是6.4节的格拉姆－施密特方法). 由于这个本征空间是2维的(有基 $\boldsymbol{v}_1,\boldsymbol{v}_2$)，根据基本定理，正交集 $\{\boldsymbol{v}_1,\boldsymbol{z}_2\}$ 是这个空间的一组正交基(见2.9节或4.5节).

标准化 $\boldsymbol{v}_1$ 和 $\boldsymbol{z}_2$ 以后，我们得到 $\lambda=7$ 对应的本征空间的标准正交基：

$$\boldsymbol{u}_1=\begin{bmatrix}1/\sqrt{2}\\0\\1/\sqrt{2}\end{bmatrix},\quad \boldsymbol{u}_2=\begin{bmatrix}-1/\sqrt{18}\\4/\sqrt{18}\\1/\sqrt{18}\end{bmatrix}$$

$\lambda=-2$ 对应本征空间的一组标准正交基是：

$$\boldsymbol{u}_3=\frac{1}{\|2\boldsymbol{v}_3\|}2\boldsymbol{v}_3=\frac{1}{3}\begin{bmatrix}-2\\-1\\2\end{bmatrix}=\begin{bmatrix}-2/3\\-1/3\\2/3\end{bmatrix}$$

451

由定理1可知，$\boldsymbol{u}_3$ 与另两个本征向量 $\boldsymbol{u}_1$ 及 $\boldsymbol{u}_2$ 正交. 于是 $\{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3\}$ 是标准正交集. 令

$$\boldsymbol{P}=[\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3]=\begin{bmatrix}1/\sqrt{2}&-1/\sqrt{18}&-2\sqrt{3}\\0&4/\sqrt{18}&-1/\sqrt{3}\\1/\sqrt{2}&1/\sqrt{18}&2/3\end{bmatrix},\quad \boldsymbol{D}=\begin{bmatrix}7&0&0\\0&7&0\\0&0&-2\end{bmatrix}$$

则 P 正交对角化 A，并且 $A=PDP^{-1}$. ■

例题3中，本征值7重数为2，其对应本征空间维数为2. 这个事实并非偶然，下面的定理可以说明这一点.

7.1.1　谱定理

矩阵 A 的本征值所构成的集合有时称作 A 的谱，下列关于本征值的命题称为谱定理.

【定理3】　对称矩阵的谱定理

一个 $n\times n$ 对称矩阵 A 具有下面性质：

a. 若计重数，A 有 n 个实本征值.

b. 每个本征值 λ 所对应本征空间的维数，等于 λ 作为特征方程的根的重数.

c. 本征空间两两正交，即，对应于不同本征值的向量互相正交.

d. A 可正交对角化.

(a)可由5.5节习题24推知. (b)易由(d)得出(见习题31). (c)就是定理1. 已知(a)，我们可以利用习题32以及第6章补充题16中讨论的舒尔分解，给出(d)的一种证明. 具体细节此处从略.

7.1.2　谱分解

设 $A=PDP^{-1}$，其中 P 的列是 A 的标准正交本征向量 $u_1,\cdots,u_n$，对应的本征值 $\lambda_1,\cdots,\lambda_n$ 是对角阵 D 中的元素. 则由 $P^{-1}=P^{\mathrm{T}}$ 有

$$A=PDP^{\mathrm{T}}=[u_1\cdots u_n]\begin{bmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{bmatrix}\begin{bmatrix}u_1^{\mathrm{T}}\\ \vdots\\ u_n^{\mathrm{T}}\end{bmatrix}=[\lambda_1u_1\cdots\lambda_nu_n]\begin{bmatrix}u_1^{\mathrm{T}}\\ \vdots\\ u_n^{\mathrm{T}}\end{bmatrix}$$

452

利用乘法的行-列展开(2.4节定理10)，我们有

$$A=\lambda_1u_1u_1^{\mathrm{T}}+\lambda_2u_2u_2^{\mathrm{T}}+\cdots+\lambda_nu_nu_n^{\mathrm{T}} \tag{2}$$

A 的这种表示将 A 分解为由谱(本征值)确定的小块，所以称作 A 的**谱分解**(spectral decomposition). (2)中每一项都是一个秩为1的 $n\times n$ 矩阵. 例如，$\lambda_1u_1u_1^{\mathrm{T}}$ 的各列均为 u_1 的倍数. 更进一步，每个 $u_ju_j^{\mathrm{T}}$ 矩阵都是一个**投影矩阵**(projection matrix)，即，对 $x\in\mathbb{R}^n$，向量 $(u_ju_j^{\mathrm{T}})x$ 是 x 在 u_j 所张成子空间上的正交投影.

【例题4】　构造矩阵 A 的一个谱分解，其中 A 可进行如下正交对角化：

$$A=\begin{bmatrix}7 & 2\\ 2 & 4\end{bmatrix}=\begin{bmatrix}2/\sqrt{5} & -1/\sqrt{5}\\ 1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}\begin{bmatrix}8 & 0\\ 0 & 3\end{bmatrix}\begin{bmatrix}2/\sqrt{5} & 1/\sqrt{5}\\ -1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}$$

解： 以 u_1，u_2 记 P 的列. 则有

$$A=8u_1u_1^{\mathrm{T}}+3u_2u_2^{\mathrm{T}}$$

为验证 A 的这一分解，计算

$$u_1u_1^{\mathrm{T}}=\begin{bmatrix}2/\sqrt{5}\\ 1/\sqrt{5}\end{bmatrix}[2/\sqrt{5}\quad 1/\sqrt{5}]=\begin{bmatrix}4/5 & 2/5\\ 2/5 & 1/5\end{bmatrix}$$

$$u_2u_2^T=\begin{bmatrix}-1/\sqrt{5}\\2/\sqrt{5}\end{bmatrix}\begin{bmatrix}-1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}=\begin{bmatrix}1/5 & -2/5\\-2/5 & 4/5\end{bmatrix}$$

且

$$8u_1u_1^T+3u_2u_2^T=\begin{bmatrix}32/5 & 16/5\\16/5 & 8/5\end{bmatrix}+\begin{bmatrix}3/5 & -6/5\\-6/5 & 12/5\end{bmatrix}=\begin{bmatrix}7 & 2\\2 & 4\end{bmatrix}=A$$ ■

注记

当 A 是一个阶数不太大的对称矩阵时，现代高性能的计算机算法可以非常精确地计算出本征值和本征向量. 他们对 A 施行一系列包含正交矩阵的相似变换. 变换矩阵的对角线元素迅速收敛于 A 的本征值(见 5.2 节注记). 利用正交矩阵，通常可以避免这一过程中累积的数值误差. 当 A 对称时，这一系列正交矩阵组合起来构成一个新的正交矩阵，它的列就是 A 的本征向量.

非对称矩阵的本征向量不可能完全正交，但这个算法仍然可以给出相当精确的本征值，接下来，就需要用非正交方法计算本征向量.

453

基础练习

1. 证明如果 A 是对称矩阵，则 A^2 也是对称的.
2. 证明如果 A 可正交对角化，则 A^2 亦可.

习题 7.1

判断习题 1 ~ 6 中哪些矩阵是对称矩阵.

1. $\begin{bmatrix}3 & 5\\5 & -7\end{bmatrix}$　2. $\begin{bmatrix}-3 & 5\\-5 & 3\end{bmatrix}$　3. $\begin{bmatrix}2 & 2\\4 & 4\end{bmatrix}$　4. $\begin{bmatrix}0 & 8 & 3\\8 & 0 & -2\\3 & -2 & 0\end{bmatrix}$

5. $\begin{bmatrix}-6 & 2 & 0\\0 & -6 & 2\\0 & 0 & -6\end{bmatrix}$　6. $\begin{bmatrix}1 & 2 & 1 & 2\\2 & 1 & 2 & 1\\1 & 2 & 1 & 2\end{bmatrix}$

判断习题 7 ~ 12 中哪些矩阵是正交矩阵. 如果是，求其逆.

7. $\begin{bmatrix}0.6 & 0.8\\0.8 & -0.6\end{bmatrix}$　8. $\begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2}\\1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}$　9. $\begin{bmatrix}-5 & 2\\2 & 5\end{bmatrix}$　10. $\begin{bmatrix}-1 & 2 & 2\\2 & -1 & 2\\2 & 2 & -1\end{bmatrix}$

11. $\begin{bmatrix}2/3 & 2/3 & 1/3\\0 & 1/\sqrt{5} & -2/\sqrt{5}\\\sqrt{5}/3 & -4/\sqrt{45} & -2/\sqrt{45}\end{bmatrix}$　12. $\begin{bmatrix}0.5 & 0.5 & -0.5 & -0.5\\-0.5 & 0.5 & -0.5 & 0.5\\0.5 & 0.5 & 0.5 & 0.5\\-0.5 & 0.5 & 0.5 & -0.5\end{bmatrix}$

正交对角化习题 13 ~ 22 中的矩阵，给出正交矩阵 P 和对角阵 D. 为节约时间，这里给出习题 17 ~ 22 的本征值：(17)5, 2, −2；(18)25, 3, −50；(19)7, −2；(20)13, 7, 1；(21)9, 5, 1；(22)2, 0.

13. $\begin{bmatrix}3 & 1\\1 & 3\end{bmatrix}$　14. $\begin{bmatrix}1 & 5\\5 & 1\end{bmatrix}$　15. $\begin{bmatrix}16 & -4\\-4 & 1\end{bmatrix}$　16. $\begin{bmatrix}-7 & 24\\24 & 7\end{bmatrix}$　17. $\begin{bmatrix}1 & 1 & 3\\1 & 3 & 1\\3 & 1 & 1\end{bmatrix}$　18. $\begin{bmatrix}-2 & -36 & 0\\-36 & -23 & 0\\0 & 0 & 3\end{bmatrix}$

19. $\begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$　20. $\begin{bmatrix} 7 & -4 & 4 \\ -4 & 5 & 0 \\ 4 & 0 & 9 \end{bmatrix}$　21. $\begin{bmatrix} 4 & 1 & 3 & 1 \\ 1 & 4 & 1 & 3 \\ 3 & 1 & 4 & 1 \\ 1 & 3 & 1 & 4 \end{bmatrix}$　22. $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

23. 设 $A=\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$，$v=\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. 验证：2 是 A 的一个本征值，且 v 是一个本征向量. 然后正交对角化 A.

24. 设 $A=\begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & 2 \\ -2 & 2 & 2 \end{bmatrix}$，$v_1=\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$及 $v_2=\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$. 验证：$v_1$ 和 v_2 是 A 的本征向量. 然后正交对角化 A.

判断习题 25 和 26 中各命题的真假，并说明理由.

25. a. 一个可正交对角化的 $n\times n$ 矩阵一定是对称的.
    b. 如果 $A^T=A$，并且向量 u 和 v 满足 $Au=3u$ 和 $Av=4v$，则 $u\cdot v=0$.
    c. 一个 $n\times n$ 对称矩阵有 n 个不等的实本征值.
    d. 对 $\mathbf{R}^n$ 中非零向量 v，矩阵 vv^T 被称作是一个投影矩阵.
26. a. 每个对称矩阵都可正交对角化.
    b. 如果 $B=PDP^{-1}$，其中 $P^T=P^{-1}$ 且 D 是一个对角阵，则 B 是一个对称矩阵.
    c. 正交矩阵可正交对角化.
    d. 对称矩阵本征空间的维数等于对应本征值的重数.
27. 假设 A，B 分别是 $n\times n$ 和 $n\times m$ 矩阵. 证明 B^TAB，B^TB 和 BB^T 是对称矩阵. 454
28. 证明：如果 A 是一个 $n\times n$ 对称矩阵，则对任意 $x,y\in\mathbf{R}^n$，都有 $(Ax)\cdot y=x\cdot(Ay)$.
29. 假设 A 可逆并且可正交对角化. 解释 A^{-1} 为什么也可正交对角化.
30. 假设 A，B 都可正交对角化，且 $AB=BA$. 解释 AB 为什么也可正交对角化.
31. 设 $A=PDP^{-1}$，其中 P 是正交矩阵，D 是对角阵. 设 λ 是 A 的一个重数为 k 的本征值. 则 λ 在 D 的对角线上出现 k 次. 解释为什么对应本征空间的维数是 k.
32. 设 $A=PRP^{-1}$，其中 P 是正交矩阵，R 是上三角矩阵. 证明：如果 A 对称，则 R 也对称，因此 R 实际上是对角阵.
33. 构造例题 2 中矩阵 A 的一个谱分解.
34. 构造例题 3 中矩阵 A 的一个谱分解.
35. 设 u 是一个 $\mathbf{R}^n$ 中的单位向量，且令 $B=uu^T$.
    a. 任意给定 $x\in\mathbf{R}^n$，计算 Bx 并证明：如 6.2 节中所言，Bx 是 x 在 u 上的正交投影.
    b. 证明：B 是对称矩阵且 $B^2=B$.
    c. 证明 u 是 B 的一个本征向量. 其对应本征值是什么？
36. 设 B 是一个 $n\times n$ 对称矩阵，且满足 $B^2=B$. 这样的矩阵被称作**投影矩阵**(projection matrix)[或者**正交投影矩阵**(orthogonal projection matrix)]. 任意给定 $y\in\mathbf{R}^n$，令 $\hat{y}=By$，$z=y-\hat{y}$.
    a. 证明 z 与 $\hat{y}$ 正交.
    b. 设 W 是 B 的列空间. 证明 y 是 W 中某向量与 $W^\perp$ 中某向量之和. 这为什么就证明了 By 是 y 在 B 的列空间上的正交投影？

[M]正交对角化习题 37～40 中的矩阵. 为了练习本节介绍的方法，不使用矩阵软件中的本征向量指令，而是仿照例题 2 和 3，利用程序求本征值，然后对每个本征值 λ，求 Nul

$(A-\lambda I)$的一组标准正交基.

37. $\begin{bmatrix} 5 & 2 & 9 & -6 \\ 2 & 5 & -6 & 9 \\ 9 & -6 & 5 & 2 \\ -6 & 9 & 2 & 5 \end{bmatrix}$

38. $\begin{bmatrix} 0.38 & -0.18 & -0.06 & -0.04 \\ -0.18 & 0.59 & -0.04 & 0.12 \\ -0.06 & -0.04 & 0.47 & -0.12 \\ -0.04 & 0.12 & -0.12 & 0.41 \end{bmatrix}$

39. $\begin{bmatrix} 0.31 & 0.58 & 0.08 & 0.44 \\ 0.58 & -0.56 & 0.44 & -0.58 \\ 0.08 & 0.44 & 0.19 & -0.08 \\ 0.44 & -0.58 & -0.08 & 0.31 \end{bmatrix}$

40. $\begin{bmatrix} 10 & 2 & 2 & -6 & 9 \\ 2 & 10 & 2 & -6 & 9 \\ 2 & 2 & 10 & -6 & 9 \\ -6 & -6 & -6 & 26 & 9 \\ 9 & 9 & 9 & 9 & -19 \end{bmatrix}$

基础练习答案

1. 根据转置矩阵的性质有$(A^2)^T=(AA)^T=A^TA^T$. 由假设可知$A^T=A$. 所以$(A^2)^T=AA=A^2$，这就说明A^2是对称的.
2. 如果A可正交对角化，则由定理2知A是对称的. 根据基础练习1，A^2是对称的，因此也是可正交对角化的(定理2).

7.2　二　次　型

到目前为止，本书中除第6章计算x^Tx时提到的平方和以外，我们的注意力都集中在线性方程上面. 将线性代数应用于工程(在设计标准和优化方面)和信号处理(输出噪声功率)时，常常出现这类和式及称为二次型的许多更一般的表达式. 它们在其他地方也有出现，比如物理(势能和动能)、微分几何(曲面的法曲率)、经济学(效用函数)和统计学(置信椭圆). 通过研究对称矩阵，我们很容易掌握这些实际问题的数学背景.

455

$\mathbf{R}^n$中的一个**二次型**(quadratic form)是定义在$\mathbf{R}^n$上的一个函数Q，它在$\mathbf{R}^n$中向量x处的值可用$Q(x)=x^TAx$形式的表达式来计算，其中A是一个$n\times n$对称矩阵. 称矩阵A为**二次型的矩阵**(matrix of the quadratic form).

非零二次型的一个最简单的例子是$Q(x)=x^TIx=\|x\|^2$. 例题1和2说明了对称矩阵A与二次型x^TAx之间的联系.

【例题1】　设$x=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. 对下列矩阵计算$x^TAx$.

a. $A=\begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$　　b. $A=\begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$

解：

a. $x^TAx=[x_1\quad x_2]\begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=[x_1\quad x_2]\begin{bmatrix} 4x_1 \\ 3x_2 \end{bmatrix}=4x_1^2+3x_2^2.$

b. A中有两个值为-2的元素. 观察它们是如何参与下述运算的，其中A的$(1,2)$-元以粗体标记.

$$\begin{aligned}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} &= [x_1 \quad x_2]\begin{bmatrix} 3 & -\mathbf{2} \\ -2 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [x_1 \quad x_2]\begin{bmatrix} 3x_1 - \mathbf{2}x_2 \\ -2x_1 + 7x_2 \end{bmatrix} \\ &= x_1(3x_1 - \mathbf{2}x_2) + x_2(-2x_1 + 7x_2) \\ &= 3x_1^2 - \mathbf{2}x_1x_2 - 2x_2x_1 + 7x_2^2 \\ &= 3x_1^2 - 4x_1x_2 + 7x_2^2 \end{aligned}$$ ■

例题 1b 二次型中 $-4x_1x_2$ 项的出现是由于矩阵 $\boldsymbol{A}$ 对角线以外有元素 -2. 与之相反，例题 1a 对角阵 $\boldsymbol{A}$ 所关联的二次型中没有 $\boldsymbol{x}_1\boldsymbol{x}_2$ 这一交叉乘积项.

【例题 2】 对于 $\boldsymbol{x} \in \mathbf{R}^n$，令 $\boldsymbol{Q}(\boldsymbol{x}) = 5\boldsymbol{x}_1^2 + 3\boldsymbol{x}_2^2 + 2\boldsymbol{x}_3^2 - \boldsymbol{x}_1\boldsymbol{x}_2 + 8\boldsymbol{x}_2\boldsymbol{x}_3$. 试将这个二次型写成 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的形式.

解： x_1^2, x_2^2 和 x_3^2 的系数位于 $\boldsymbol{A}$ 的对角线上. 为使 $\boldsymbol{A}$ 对称，$x_ix_j(i \neq j)$ 的系数必须平分给 $\boldsymbol{A}$ 的 (i,j) 元和 (j,i) 元. x_1x_3 的系数是 0. 容易验证

$$\boldsymbol{Q}(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = [x_1 \quad x_2 \quad x_3]\begin{bmatrix} 5 & -1/2 & 0 \\ -1/2 & 3 & 4 \\ 0 & 4 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$ ■

【例题 3】 设 $\boldsymbol{Q}(\boldsymbol{x}) = x_1^2 - 8x_1x_2 - 5x_2^2$. 对 $\boldsymbol{x} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \end{bmatrix}$ 和 $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ 分别计算 $\boldsymbol{Q}(\boldsymbol{x})$ 的值. 456

解：

$$\begin{aligned} Q(-3,1) &= (-3)^2 - 8(-3)(1) - 5(1)^2 = 28 \\ Q(2,-2) &= (2)^2 - 8(2)(-2) - 5(-2)^2 = 16 \\ Q(1,-3) &= (1)^2 - 8(1)(-3) - 5(-3)^2 = -20 \end{aligned}$$ ■

在某些情形下，当二次型不含交叉乘积项，即，当二次型的矩阵是对角阵时，二次型更易于应用. 庆幸的是，交叉乘积项可以通过适当的变量替换来消去.

7.2.1 二次型中的变量替换

若 $\boldsymbol{x}$ 代表 $\mathbf{R}^n$ 中的一个可变向量，则一个**变量替换**(change of variable)是一个形如

$$\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y}，\text{或等价地，} \boldsymbol{y} = \boldsymbol{P}^{-1}\boldsymbol{x} \tag{1}$$

的方程，其中 $\boldsymbol{P}$ 是可逆矩阵，$\boldsymbol{y}$ 是 $\mathbf{R}^n$ 中一个新的可变向量. 这里 $\boldsymbol{y}$ 是 $\boldsymbol{x}$ 在 $\boldsymbol{P}$ 的列所确定的 $\mathbf{R}^n$ 的那组基之下的坐标向量(见 4.4 节).

如果在二次型中施以变量替换(1)，则有

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{P}\boldsymbol{y})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{P}\boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P})\boldsymbol{y} \tag{2}$$

且这个二次型的新矩阵是 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$. 如果 $\boldsymbol{P}$ 正交对角化 $\boldsymbol{A}$，则 $\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{P}^{-1}$ 且 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D}$. 于是这个二次型的矩阵就是对角阵！下面的例题就将使用这个策略.

【例题 4】 做一个变量替换，将例题 3 的二次型变成不含交叉乘积项的二次型.

解： 例题 3 中二次型的矩阵是：

$$\boldsymbol{A} = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

第一步是正交对角化 $\boldsymbol{A}$. 求出其本征值为 $\lambda=3$ 和 $\lambda=-7$. 对应的单位本征向量为：

$$\lambda=3:\ \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}; \qquad \lambda=-7:\ \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

这两个向量自动正交(因为它们对应于不同的本征值)，所以构成 $\mathbf{R}^2$ 的一组标准正交基. 令

$$\boldsymbol{P}=\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}, \qquad \boldsymbol{D}=\begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$

457 则如前文所述，有 $\boldsymbol{A}=\boldsymbol{PDP}^{-1}$ 及 $\boldsymbol{D}=\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$. 所求变量替换为

$$\boldsymbol{x}=\boldsymbol{Py}, \qquad \text{其中 } \boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \boldsymbol{y}=\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

则

$$x_1^2-8x_1x_2-5x_2^2=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}=(\boldsymbol{Py})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{Py})=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{APy}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{Dy}=3y_1^2-7y_2^2$$ ■

为了图解例题4中二次型相等的含义，我们利用新的二次型，计算 $\boldsymbol{x}=(2,-2)$ 时 $\boldsymbol{Q}(\boldsymbol{x})$ 的值. 首先，由于 $\boldsymbol{x}=\boldsymbol{Py}$，我们有

$$\boldsymbol{y}=\boldsymbol{P}^{-1}\boldsymbol{x}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{x}$$

所以

$$\boldsymbol{y}=\begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}\begin{bmatrix} 2 \\ -2 \end{bmatrix}=\begin{bmatrix} 6/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

于是

$$3y_1^2-7y_2^2=3(6/\sqrt{5})^2-7(-2/\sqrt{5})^2=3(36/5)-7(4/5)=80/5=16$$

这就是例题3中 $\boldsymbol{Q}(\boldsymbol{x})$ 在 $\boldsymbol{x}=(2,-2)$ 时的值. 见图7-2.

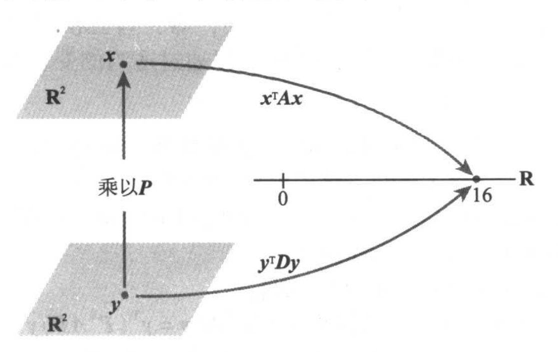


图7-2 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}$ 中的变量替换

例题4实际阐明了下面的定理. 其证明早在例题4之前就给出了.

【定理4】 主轴定理

设 $\boldsymbol{A}$ 是一个 $n\times n$ 对称矩阵. 则存在一个变量的正交变换 $\boldsymbol{x}=\boldsymbol{Py}$，它将二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}$ 变成不含交叉乘积项的二次型 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{Dy}$.

458

定理中 $\boldsymbol{P}$ 的列称为二次型的**主轴**(principal axis). 向量 $\boldsymbol{y}$ 是 $\boldsymbol{x}$ 在主轴所构成的 $\mathbf{R}^n$ 的标准正交基之下的坐标向量.

7.2.2 主轴的几何意义

设 $Q(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$, 其中 $\boldsymbol{A}$ 是一个 2×2 可逆对称矩阵. 令 c 为一常数. 可以证明 $\mathbf{R}^n$ 中满足

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=c \tag{3}$$

的全体 $\boldsymbol{x}$ 所构成的集合或者是一个椭圆(或圆)、双曲线、两条相交直线, 或者是一个单点, 或者是空集. 若 $\boldsymbol{A}$ 是对角阵, 则图像位于标准位置, 如图 7-3 所示. 若 $\boldsymbol{A}$ 不是对角阵, 则(3)的图像旋转到标准位置以外, 如图 7-4. 求($\boldsymbol{A}$ 的本征向量所确定的)主轴就相当于求使图像位于标准位置的一个新坐标系.

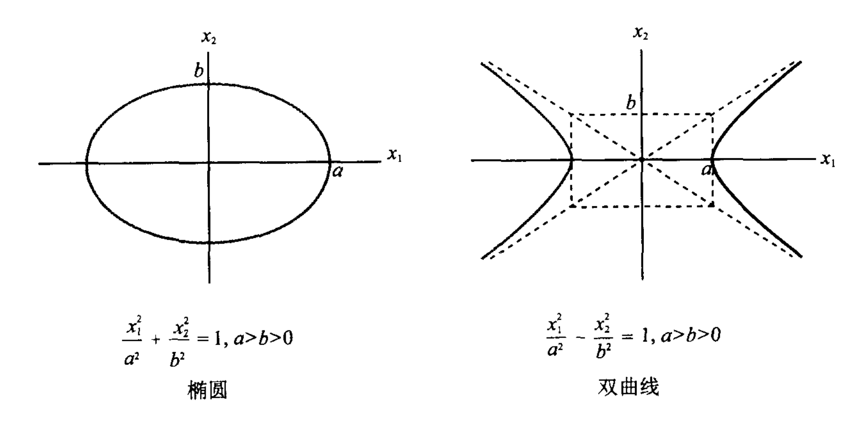


图 7-3 标准位置上的椭圆和双曲线

图 7-4b 中的双曲线是方程 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=16$ 的图像, 其中 $\boldsymbol{A}$ 是例题 4 中的矩阵. 图7-4b y_1 轴正向与例题 4 中 $\boldsymbol{P}$ 的首列同向, y_2 轴正向与 $\boldsymbol{P}$ 的第二列同向.

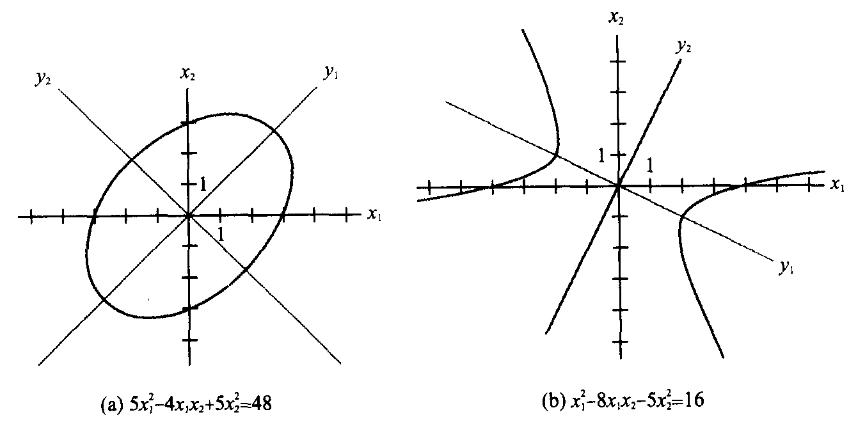


图 7-4 不在标准位置上的椭圆和双曲线

【例题 5】　图 7-4a 中的椭圆是方程 $5x_1^2-4x_1x_2+5x_2^2=48$ 的图像. 求能够消去方程中交叉乘积项的一个变量替换.

解： 二次型的矩阵是 $A=\begin{bmatrix}5 & -2\\ -2 & 5\end{bmatrix}$. 求出 A 的本征值为 3 和 7，其对应单位本征向量是：

459

$$u_1=\begin{bmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix},\quad u_2=\begin{bmatrix}-1\ \sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}$$

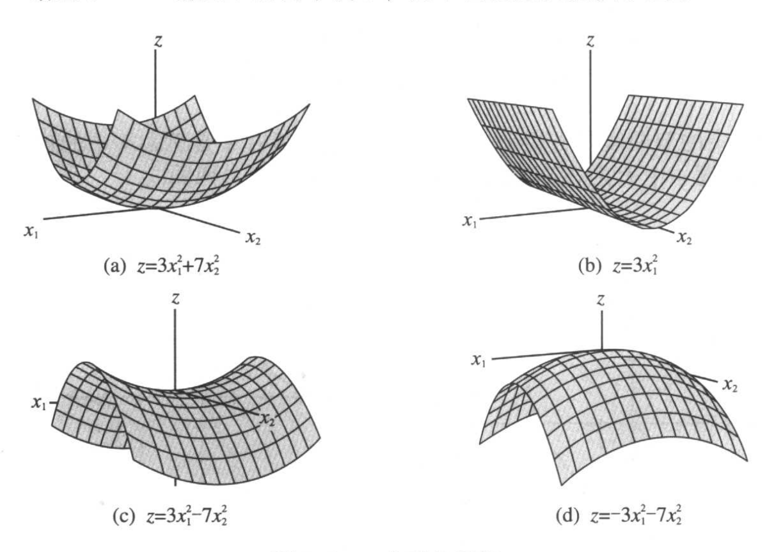
令 $P=[u_1\quad u_2]=\begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2}\\ 1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}$. 则 P 正交对角化 A，所以变量替换 $x=Py$ 导出二次型 $y^TDy=3y_1^2+7y_2^2$. 这个变量替换的新主轴如图 7-4a 所示. ■

7.2.3　二次型的分类

当 A 是一个 $n\times n$ 矩阵时，二次型 $Q(x)=x^TAx$ 是定义域为 $\mathbf{R}^n$ 的一个实值函数. 下面我们根据二次型在不同 x 处取值的类型，将它们划分为重要的几大类.

图 7-5 给出了四个二次型的图像. 对于二次型定义域中的每个点 $x=(x_1,x_2)$，图中绘出了点 (x_1,x_2,z)，其中 $z=Q(x)$. 注意除 $x=0$ 以外，图 7-5a 中 $Q(x)$ 取值总为正，而图 7-5d 中总为负. 图 7-5a 和图 7-5d 的水平横截面是椭圆，而图7-5c的是双曲线.

图 7-5 给出了 2×2 情况下的简单例子，为下面的定义提供了图示.



(a) $z=3x_1^2+7x_2^2$　(b) $z=3x_1^2$　(c) $z=3x_1^2-7x_2^2$　(d) $z=-3x_1^2-7x_2^2$

460

图 7-5　二次型的图像

【定义】　称一个二次型 Q：

a. **正定**(positive definite)，如果对任意 $x\neq0$ 都有 $Q(x)>0$；

b. **负定**(negative definite)，如果对任意 $\boldsymbol{x} \neq \boldsymbol{0}$ 都有 $Q(\boldsymbol{x})<0$；
c. **不定**(indefinite)，如果 $Q(\boldsymbol{x})$ 既可取正值也可取负值.

此外，称 Q **半正定**(positive semidefinite)，如果对任意 $\boldsymbol{x}$ 都有 $Q(\boldsymbol{x}) \geqslant 0$，称 Q **半负定**(negative semidefinite)，如果对任意 $\boldsymbol{x}$ 都有 $Q(\boldsymbol{x}) \leqslant 0$. 图 7-5(a)和(b)中的二次型都是半正定的.

【定理 5】　二次型与本征值

设 $\boldsymbol{A}$ 是一个 $n \times n$ 对称矩阵. 则二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$：
a. 是正定的当且仅当 $\boldsymbol{A}$ 的本征值都是正的；
b. 是负定的当且仅当 $\boldsymbol{A}$ 的本征值都是负的；
c. 是不定的当且仅当 $\boldsymbol{A}$ 的本征值既有正也有负.

证明：根据主轴定理，存在变量替换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 使得：

$$Q(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{y}=\lambda_1 y_1^2+\lambda_2 y_2^2+\cdots+\lambda_n y_n^2 \tag{4}$$

其中 λ_1，…，λ_n 是 $\boldsymbol{A}$ 的本征值. 由于 $\boldsymbol{P}$ 可逆，全体非零的 $\boldsymbol{x}$ 到全体非零的 $\boldsymbol{y}$ 之间存在一个一一对应. 于是 $\boldsymbol{x} \neq \boldsymbol{0}$ 时 $Q(\boldsymbol{x})$ 的值与(4)右端表达式的值一致，而这个值显然由 λ_1，…，λ_n 的符号所控制，并且符合定理所述的三条. ■

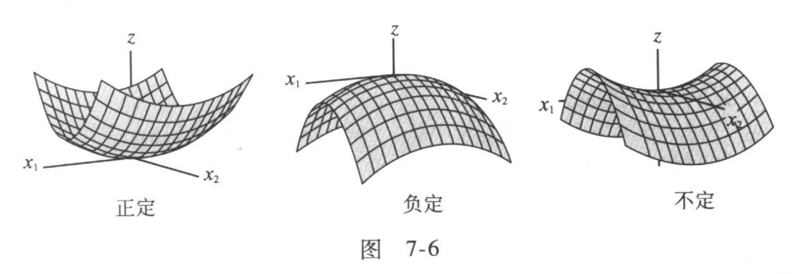


图 7-6

【例题 6】　$Q(\boldsymbol{x})=3x_1^2+2x_2^2+x_3^2+4x_1x_2+4x_2x_3$ 是否为正定的？

解：由于所出现的符号都是加号，这个二次型看似正定. 但其矩阵是

$$\boldsymbol{A}=\begin{bmatrix}3 & 2 & 0\\ 2 & 2 & 2\\ 0 & 2 & 1\end{bmatrix}$$

且算出 $\boldsymbol{A}$ 的本征值为 5，2 和 −1. 所以 Q 是不定二次型，而不是正定二次型. ■

二次型的分类常常被拓展到相应矩阵的分类上. 因此，一个**正定矩阵**(positive definite matrix)就是一个使二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定的对称矩阵 $\boldsymbol{A}$. 其他的概念，比如**半正定矩阵**(positive semidefinite matrix)，也有类似的定义. 461

注记

判断对称矩阵 $\boldsymbol{A}$ 是否正定的一个快速方法是尝试将 $\boldsymbol{A}$ 分解成 $\boldsymbol{A}=\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}$ 的形式，其中 $\boldsymbol{R}$ 是对角线元素为正的上三角矩阵(一种途径是对 LU 分解的算法稍加修改). 这样的一个楚列斯基分解是可行的当且仅当 $\boldsymbol{A}$ 正定. 见补充题 7.

基础练习

试用本征值来描述半正定矩阵 $\boldsymbol{A}$.

习题 7.2

1. 计算二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$，其中 $\boldsymbol{A}=\begin{bmatrix}5 & 1/3\\ 1/3 & 1\end{bmatrix}$，且

   a. $\boldsymbol{x}=\begin{bmatrix}x_1\\ x_2\end{bmatrix}$ b. $\boldsymbol{x}=\begin{bmatrix}6\\ 1\end{bmatrix}$ c. $\boldsymbol{x}=\begin{bmatrix}1\\ 3\end{bmatrix}$

2. 计算二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$，其中 $\boldsymbol{A}=\begin{bmatrix}4 & 3 & 0\\ 3 & 2 & 1\\ 0 & 1 & 1\end{bmatrix}$，且

   a. $x=\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$ b. $\boldsymbol{x}=\begin{bmatrix}2\\ -1\\ 5\end{bmatrix}$ c. $\boldsymbol{x}=\begin{bmatrix}1/\sqrt{3}\\ 1/\sqrt{3}\\ 1/\sqrt{3}\end{bmatrix}$

3. 设 $\boldsymbol{x}\in\mathbf{R}^2$，求下列二次型的矩阵：
   a. $10x_1^2-6x_1x_2-3x_2^2$ b. $5x_1^2+3x_1x_2$
4. 设 $\boldsymbol{x}\in\mathbf{R}^2$，求下列二次型的矩阵：
   a. $20x_1^2+15x_1x_2-10x_2^2$ b. x_1x_2
5. 设 $\boldsymbol{x}\in\mathbf{R}^3$，求下列二次型的矩阵：
   a. $8x_1^2+7x_2^2-3x_3^2-6x_1x_2+4x_1x_3-2x_2x_3$ b. $4x_1x_2+6x_1x_3-8x_2x_3$
6. 设 $\boldsymbol{x}\in\mathbf{R}^3$，求下列二次型的矩阵：
   a. $5x_1^2-x_2^2+7x_3^2+5x_1x_2-3x_1x_3$ b. $x_3^2-4x_1x_2+4x_2x_3$
7. 做一个变量替换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$，将二次型 $x_1^2+10x_1x_2+x_2^2$ 变成不含交叉乘积项的二次型．写出 $\boldsymbol{P}$ 和新的二次型．
8. 设 $\boldsymbol{A}$ 是二次型 $9x_1^2+7x_2^2+11x_3^2-8x_1x_2+8x_1x_3$ 的矩阵．可以求出 $\boldsymbol{A}$ 的本征值为3,9和15．求正交矩阵 $\boldsymbol{P}$，使得变量替换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 将 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 变成不含交叉乘积项的二次型．写出 $\boldsymbol{P}$ 和新的二次型．

   对习题9～18中的二次型分类．然后做变量替换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$，将它们变成不含交叉乘积项的二次型．并写出新的二次型．可以利用7.1节的方法构造 $\boldsymbol{P}$．

9. $3x_1^2-4x_1x_2+6x_2^2$
10. $9x_1^2-8x_1x_2+3x_2^2$
11. $2x_1^2+10x_1x_2+2x_2^2$
12. $-5x_1^2+4x_1x_2-2x_2^2$
13. $x_1^2-6x_1x_2+9x_2^2$
14. $8x_1^2+6x_1x_2$
15. [M] $-2x_1^2-6x_2^2-9x_3^2-9x_4^2+4x_1x_2+4x_1x_3+4x_1x_4+6x_3x_4$
16. [M] $4x_1^2+4x_2^2+4x_3^2+4x_4^2+3x_1x_2+3x_3x_4-4x_1x_4+4x_2x_3$
17. [M] $x_1^2+x_2^2+x_3^2+x_4^2+9x_1x_2-12x_1x_4+12x_2x_3+9x_3x_4$
18. [M] $11x_1^2-x_2^2-12x_1x_2-12x_1x_3-12x_1x_4-2x_3x_4$
19. 如果 $\boldsymbol{x}=(x_1,x_2)$ 且 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$，即 $x_1^2+x_2^2=1$，则二次型 $5x_1^2+8x_2^2$ 可能取到的最大值是多少？（对 $\boldsymbol{x}$ 举几个例子试试．）
20. 如果 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$，则二次型 $5x_1^2-3x_2^2$ 可能取到的最大值是多少？

462 习题21和22中，矩阵是 $n\times n$ 阵，向量属于 $\mathbf{R}^n$．判断各命题的真假，并说明理由．

21. a. 二次型的矩阵是对称矩阵．
    b. 一个二次型不含交叉乘积项当且仅当其对应矩阵是对角阵．

c. 二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的主轴是 $\boldsymbol{A}$ 的本征向量.

d. 一个正定二次型 $\boldsymbol{Q}$ 满足：对任意 $\boldsymbol{x}\in\mathbf{R}^n$，有 $\boldsymbol{Q}(\boldsymbol{x})>0$.

e. 如果对称矩阵 $\boldsymbol{A}$ 的本征值都是正的，则二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定.

f. 对称矩阵的一个楚列斯基分解形如 $\boldsymbol{A}=\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}$，其中 $\boldsymbol{R}$ 是对角线元素为正的上三角矩阵.

22. a. 表达式 $\|\boldsymbol{x}\|^2$ 是一个二次型.

b. 如果 $\boldsymbol{A}$ 是对称矩阵且 $\boldsymbol{P}$ 是正交矩阵，则变量替换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 将 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 变成不含交叉乘积项的二次型.

c. 如果 $\boldsymbol{A}$ 是一个 2×2 对称矩阵，则满足 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=c$（$c$ 为某一常数）的 $\boldsymbol{x}$ 所构成的集合对应于一个圆，一个椭圆或者一条双曲线.

d. 一个不定二次型或者半正定，或者半负定.

e. 如果 $\boldsymbol{A}$ 是对称矩阵，并且 $\boldsymbol{x}\neq\boldsymbol{0}$ 时二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 只取负值，则 $\boldsymbol{A}$ 的本征值都是负的.

习题23和24演示了：不计算 $\boldsymbol{A}$ 的本征值，怎样对二次型 $\boldsymbol{Q}(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 进行分类. 其中 $\boldsymbol{A}=\begin{bmatrix} a & b \\ b & d \end{bmatrix}$ 且 $\det\boldsymbol{A}\neq0$.

23. 如果 λ_1 和 λ_2 是 $\boldsymbol{A}$ 的本征值，则 $\boldsymbol{A}$ 的特征多项式可以用下面两种方式书写：$\det(\boldsymbol{A}-\lambda\boldsymbol{I})$ 和 $(\lambda-\lambda_1)(\lambda-\lambda_2)$. 利用这一事实证明：$\lambda_1+\lambda_2=a+d$（$\boldsymbol{A}$ 的对角线元素）以及 $\lambda_1\lambda_2=\det\boldsymbol{A}$.

24. 验证下列各命题：

a. 如果 $\det\boldsymbol{A}>0$ 且 $a>0$，则 $\boldsymbol{Q}$ 正定.

b. 如果 $\det\boldsymbol{A}>0$ 且 $a<0$，则 $\boldsymbol{Q}$ 负定.

c. 如果 $\det\boldsymbol{A}<0$，则 $\boldsymbol{Q}$ 不定.

25. 证明：如果 $\boldsymbol{B}$ 是 $m\times n$ 矩阵，则 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}$ 半正定；如果 $\boldsymbol{B}$ 是 $n\times n$ 可逆矩阵，则 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}$ 正定.

26. 证明：如果 $n\times n$ 矩阵 $\boldsymbol{A}$ 是正定的，则存在正定矩阵 $\boldsymbol{B}$，使得 $\boldsymbol{A}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}$. [提示：记 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{\mathrm{T}}$，其中 $\boldsymbol{P}^{\mathrm{T}}=\boldsymbol{P}^{-1}$. 构造对角阵 $\boldsymbol{C}$，使得 $\boldsymbol{D}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$，并且令 $\boldsymbol{B}=\boldsymbol{P}\boldsymbol{C}\boldsymbol{P}^{\mathrm{T}}$. 证明 $\boldsymbol{B}$ 满足命题要求.]

27. 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是 $n\times n$ 对称矩阵，且本征值都为正值. 证明 $\boldsymbol{A}+\boldsymbol{B}$ 的本征值也都是正的. [提示：考虑二次型.]

28. 设 $\boldsymbol{A}$ 是一个 $n\times n$ 可逆对称矩阵. 证明：若二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定，则二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{x}$ 亦然. [提示：考虑本征值.]

基础练习答案

做变量替换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$，与(4)中一样，记

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{y}=\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2$$

如果某个本征值，比方说 λ_i，是负的，则当 $\boldsymbol{x}$ 取 $\boldsymbol{y}=\boldsymbol{e}_i$（$\boldsymbol{I}_n$ 的第 i 列）的对应值时，$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 将取负值. 所以半正定二次型的本征值一定都非负. 反过来，如果本征值非负，上面的展开式表明 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 一定是半正定的（见图7-7）.

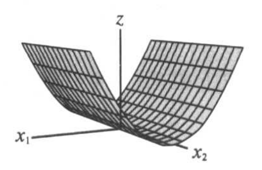


图7-7 半正定

7.3 约束优化

求二次型 $\boldsymbol{Q}(\boldsymbol{x})$ 对于指定集合中 $\boldsymbol{x}$ 的最大或最小值，是工程学家、经济学家、科学家和数学家们常常会遇到的问题. 一个典型的问题是允许 $\boldsymbol{x}$ 取遍全体单位向量. 我们将要看到，这类约束优化问题的解十分有趣且优美. 下面的例题6以及7.5节中的讨论都以实例说明了这类问题是怎样出现在实际应用中的.

463

$\mathbf{R}^n$中的向量 $\boldsymbol{x}$ 是一个单位向量，这一要求可用下列几种方式表示：

$$\|\boldsymbol{x}\|=1,\|\boldsymbol{x}\|^2=1,\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$$

或

$$x_1^2+x_2^2+\cdots+x_n^2=1 \tag{1}$$

我们将采用 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 的表示，不过，展开式(1)在应用中较为普遍.

如果二次型 $\boldsymbol{Q}$ 不含交叉乘积项，容易求得 $\boldsymbol{Q}(\boldsymbol{x})$ 在 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 条件下的最大值和最小值.

【例题 1】　求 $\boldsymbol{Q}(\boldsymbol{x})=9x_1^2+4x_2^2+3x_3^2$ 在约束条件 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 下的最大值和最小值.

解： 由于 x_2^2 和 x_3^2 非负，注意到

$$4x_2^2\leqslant 9x_2^2,,\ \text{且}\ 3x_3^2\leqslant 9x_3^2$$

于是，$x_1^2+x_2^2+x_3^2=1$ 时有：

$$\boldsymbol{Q}(\boldsymbol{x})=9x_1^2+4x_2^2+3x_3^2\leqslant 9x_1^2+9x_2^2+9x_3^2=9(x_1^2+x_2^2+x_3^2)=9$$

所以当 $\boldsymbol{x}$ 是单位向量时，$\boldsymbol{Q}(\boldsymbol{x})$的最大值不大于9. 更进一步，$\boldsymbol{x}=(1,0,0)$时有 $\boldsymbol{Q}(\boldsymbol{x})=9$. 这样，9 就是 $\boldsymbol{Q}(\boldsymbol{x})$在 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 条件下的最大值.

为了求出 $\boldsymbol{Q}(\boldsymbol{x})$的最小值，观察

$$9x_1^2\geqslant 3x_1^2,\qquad 4x_2^2\geqslant 3x_2^2$$

于是，$x_1^2+x_2^2+x_3^2=1$ 时有

$$\boldsymbol{Q}(\boldsymbol{x})\geqslant 3x_1^2+3x_2^2+3x_3^2=3(x_1^2+x_2^2+x_3^2)=3$$

此外，$x_1=0,x_2=0$ 且 $x_3=1$ 时有 $\boldsymbol{Q}(\boldsymbol{x})=3$. 所以 3 是 $\boldsymbol{Q}(\boldsymbol{x})$在 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 条件下的最小值. ■

例题 1 中，易见二次型 $\boldsymbol{Q}$ 对应的矩阵有本征值 9,4 和 3，并且最大、最小本征值分别等于(约束条件下)$\boldsymbol{Q}(\boldsymbol{x})$的最大值和最小值. 下面我们将看到，这一事实对任意二次型都成立.

464 **【例题 2】**　设 $\boldsymbol{A}=\begin{bmatrix}3&0\\0&7\end{bmatrix}$，且令 $\boldsymbol{Q}(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$，其中 $\boldsymbol{x}\in\mathbf{R}^n$. 图 7-8 绘出了 $\boldsymbol{Q}$ 的图像. 图 7-9 只显示了圆柱体内的部分图像；曲面与圆柱的交是满足 $z=\boldsymbol{Q}(x_1,x_2)$ 及 $x_1^2+x_2^2=1$ 的全体点(x_1,x_2,z)的集合. 这些点的“高度”就是在约束条件下的值. 从几何上来看，约束优化问题就是确定交线上的最高点和最低点.

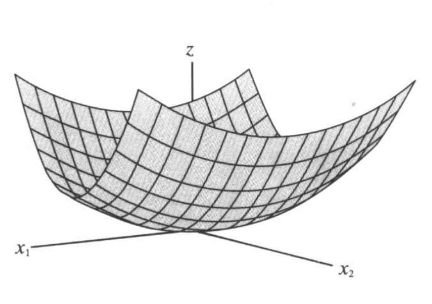


图 7-8　$z=3x_1^2+7x_2^2$

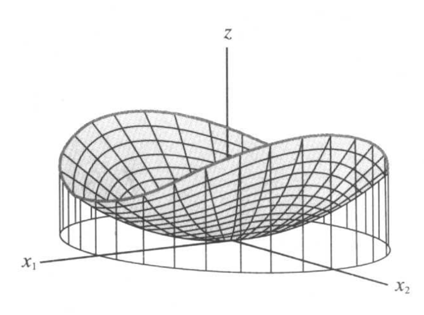


图 7-9　$z=3x_1^2+7x_2^2$ 与圆柱 $x_1^2+x_2^2=1$ 的交

这条曲线上的两个最高点出现在 $x_1=0, x_2=\pm1$ 的位置，且位于 x_1x_2 平面上方 7 个单位. 这两点对应于 $\boldsymbol{A}$ 的本征值 7，以及本征向量 $\boldsymbol{x}=(0,1)$ 和 $-\boldsymbol{x}=(0,-1)$. 类似地，曲线上的两个最低点位于 x_1x_2 平面上方 3 个单位. 它们对应于 $\boldsymbol{A}$ 的本征值 3，以及本征向量 $(1,0)$ 和 $(-1,0)$. ■

图 7-9 交线上每个点的 z 坐标都介于 3 到 7 之间，并且对 3 到 7 之间的任意数 t，都存在一个单位向量使 $Q(\boldsymbol{x})=t$. 换言之，约束条件 $\|\boldsymbol{x}\|=1$ 下，$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的所有可能的取值构成闭区间 $3\leqslant t\leqslant 7$.

可以证明，对任意对称矩阵，约束条件 $\|\boldsymbol{x}\|=1$ 下，$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的所有可能的取值都构成实数轴上的一个闭区间(见习题 13). 分别以 m 和 M 记这个区间的左右端点. 即，令

$$m=\min\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}:\ \|\boldsymbol{x}\|=1\}, M=\max\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}:\ \|\boldsymbol{x}\|=1\} \tag{2}$$

习题 12 要求证明：如果 λ 是 $\boldsymbol{A}$ 的本征值，则 $m\leqslant\lambda\leqslant M$. 下面的定理断言 m 和 M 本身就是 $\boldsymbol{A}$ 的本征值，正如例题 2 所示.[1]

【定理 6】 设 $\boldsymbol{A}$ 是对称矩阵，定义 m 和 M 如(2)式. 则 M 是 $\boldsymbol{A}$ 的最大的本征值 λ_1，m 是 $\boldsymbol{A}$ 的最小的本征值. 当 $\boldsymbol{x}$ 是与 M 对应的单位本征向量 $\boldsymbol{u}_1$ 时，$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的值是 M. 当 $\boldsymbol{x}$ 是与 m 对应的单位本征向量时，$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的值是 m.

465

证明： 设 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$ 是 $\boldsymbol{A}$ 的正交对角化形式. 我们知道，

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{y},\quad \text{当 } \boldsymbol{x}=\boldsymbol{P}\boldsymbol{y} \text{ 时} \tag{3}$$

此外，因为 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}=\boldsymbol{I}$ 以及 $\|\boldsymbol{P}\boldsymbol{y}\|^2=(\boldsymbol{P}\boldsymbol{y})^{\mathrm{T}}(\boldsymbol{P}\boldsymbol{y})=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{y}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}=\|\boldsymbol{y}\|^2$，我们还有

$$\|\boldsymbol{x}\|=\|\boldsymbol{P}\boldsymbol{y}\|=\|\boldsymbol{y}\|, \text{对任意 } \boldsymbol{y}$$

特别地，$\|\boldsymbol{y}\|=1$ 当且仅当 $\|\boldsymbol{x}\|=1$. 这样，当 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 取遍全体单位向量时，$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 和 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{y}$ 取值相同.

为简化记号，我们设 $\boldsymbol{A}$ 是 3×3 矩阵，其本征值 $a\geqslant b\geqslant c$. 适当排列 $\boldsymbol{P}$ 的(本征向量)列，使 $\boldsymbol{P}=[\boldsymbol{u}_1\quad \boldsymbol{u}_2\quad \boldsymbol{u}_3]$，且

$$\boldsymbol{D}=\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

给定 $\mathbb{R}^n$ 中单位向量 $\boldsymbol{y}$ 的坐标 y_1, y_2 和 y_3，观察

$$ay_1^2=ay_1^2$$
$$by_2^2\leqslant ay_2^2$$
$$cy_3^2\leqslant ay_3^2$$

将这些不等式累加起来，我们得到

$$\boldsymbol{y}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{y}=ay_1^2+by_2^2+cy_3^2\leqslant ay_1^2+ay_2^2+ay_3^2=a(y_1^2+y_2^2+y_3^2)=a\|\boldsymbol{y}\|^2=a$$

于是，根据 M 的定义知 $M\leqslant a$. 然而，当 $\boldsymbol{y}=\boldsymbol{e}_1=(1,0,0)$ 时 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{y}=a$，所以事实上 $M=a$. 根据(3)，与 $\boldsymbol{y}=\boldsymbol{e}_1$ 对应的 $\boldsymbol{x}$ 是 $\boldsymbol{A}$ 的本征向量 $\boldsymbol{u}_1$，这是因为

---

1. (2) 中的最大、最小，以及定理中的最大、最小都相对实数的自然顺序而言，并不是指绝对值.

$$x = Pe_1 = [u_1 \quad u_2 \quad u_3]\begin{bmatrix}1\\0\\0\end{bmatrix} = u_1$$

这样就有 $M = a = e_1^{\mathrm{T}}De_1 = u_1^{\mathrm{T}}Au_1$，这就证明了关于 M 的论断. 类似的讨论可以证明 m 就是最小的本征值 c，并且，$x = Pe_3 = u_3$ 时 $x^{\mathrm{T}}Ax$ 可以取到这个值. ■

【例题3】 设 $A = \begin{bmatrix}3&2&1\\2&3&1\\1&1&4\end{bmatrix}$. 求二次型 $x^{\mathrm{T}}Ax$ 在约束条件 $x^{\mathrm{T}}x = 1$ 下的最大值，以及

466 使 $x^{\mathrm{T}}Ax$ 取到最大值的一个单位向量.

解： 根据定理6，我们求 A 的最大本征值. 算出其特征方程为

$$0 = -\lambda^3 + 10\lambda^2 - 27\lambda + 18 = -(\lambda - 6)(\lambda - 3)(\lambda - 1)$$

最大本征值是6.

当 x 是 $\lambda = 6$ 所对应的单位本征向量时，$x^{\mathrm{T}}Ax$ 取到约束条件下的最大值. 解 $(A - 6I)x = 0$，我们求得本征向量 $\begin{bmatrix}1\\1\\1\end{bmatrix}$ 及 $u_1 = \begin{bmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$. ■

在后面的应用中，我们将进一步考虑 $x^{\mathrm{T}}Ax$ 的取值，其中 x 不但是单位向量，而且还正交于定理6中本征向量 u_1. 下面定理所讨论的就是这种情形.

【定理7】 设 A、λ_1 和 u_1 同定理6中. 则 $x^{\mathrm{T}}Ax$ 在下列约束条件下

$$x^{\mathrm{T}}x = 1, x^{\mathrm{T}}u_1 = 0$$

最大值是第二大的本征值 λ_2，并且，当 x 是 λ_2 对应的本征向量 u_2 时，$x^{\mathrm{T}}Ax$ 可以取到这个值.

通过与前面类似的讨论，我们也可以证明定理7，此时要将定理7简化成二次型矩阵是对角阵的情形. 下面的例题给出了对角阵情况下证明定理的一个思路.

【例题4】 求 $9x_1^2 + 4x_2^2 + 3x_3^2$ 在约束条件 $x^{\mathrm{T}}x = 1$ 和 $x^{\mathrm{T}}u_1 = 0$ 下的最大值，其中 $u_1 = (1,0,0)$. 注意 u_1 是与二次型矩阵的最大本征值 $\lambda = 9$ 对应的单位本征向量.

解： 如果 x 的坐标为 x_1, x_2 和 x_3，则约束条件 $x^{\mathrm{T}}u_1 = 1$ 仅仅意味着 $x_1 = 0$. 对于一个这样的单位向量，有 $x_2^2 + x_3^2 = 1$，且

$$9x_1^2 + 4x_2^2 + 3x_3^2 = 4x_2^2 + 3x_3^2 \leqslant 4x_2^2 + 4x_3^2 = 4(x_2^2 + x_3^2) = 4$$

于是，二次型在约束条件下的最大值不超过4. 并且当 $x = (0,1,0)$ 时，二次型可以取到这个值，而这个 x 就是二次型矩阵第二大本征值所对应的一个本征向量. ■

【例题5】 设矩阵 A 同例题3，且设 u_1 是与 A 的最大本征值对应的单位本征向量. 求 $x^{\mathrm{T}}Ax$ 在下列约束条件下的最大值：

467

$$x^{\mathrm{T}}x = 1, x^{\mathrm{T}}u_1 = 0 \qquad (4)$$

解： 由例题3可知，A 的第二大本征值是 $\lambda = 3$. 解 $(A - 3I)x = 0$ 可求出一个本征向量，标准化后得到

$$u_2 = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}$$

向量u_2自动与u_1正交，因为它们对应于不同的本征值. 于是x^TAx在(4)中约束条件下的最大值是3，并且这个值在$x=u_2$时被取到. ■

下面的定理是定理7和定理6的推广，它给出了A的全部本征值的刻画. 其证明在此略过.

【定理8】 设A是一个$n\times n$对称矩阵，并且有正交对角化形式$A=PDP^{-1}$，其中D的对角线元素按$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$的顺序排列，$P$的列恰为对应的单位本征向量$u_1,\cdots,u_n$. 则对$k=2,\cdots,n$，$x^TAx$在下列约束条件下的最大值是本征值$\lambda_k$，并且在$x=u_k$时可以取到这个值.

$$x^Tx=1, x^Tu_1=0,\cdots,x^Tu_{k-1}=0$$

定理8在7.4和7.5节中将非常有用. 下面的实例只需应用定理6.

【例题6】 某县政府计划于明年修建x百里公用道路和桥梁，并且修缮y百亩公园和休闲娱乐场所. 该县必须确定如何在两项工程之间配置资源(资金、设备、劳动力等等). 如果两项工程同时实施比单独实施一项要划算，则x和y必须满足如下约束：

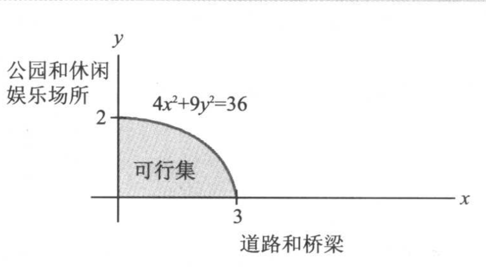


图7-10 公共工程实施计划

$$4x^2+9y^2\leqslant 36$$

见图7-10. 阴影部分代表可行集，其中任意一点(x,y)都代表了一个可行的年度公共工程实施计划. 约束曲线$4x^2+9y^2=36$上的点最充分地利用了可用的资源.

468

在确定公共工程实施计划时，该县希望考虑居民的意见. 为了度量居民们赋给不同计划(x,y)的值或*效用*，经济学家们有时会使用下列函数：

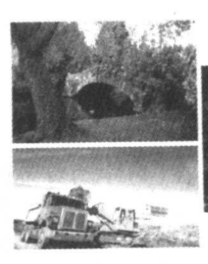
$$q(x,y)=xy$$

那些使$q(x,y)$相等的点(x,y)所构成的集合称为*无差别曲线*. 图7-11绘出了4条这样的曲线. 同一条无差别曲线上的点所对应的计划可以互相替代，居民们会发现它们具有同样的效用.[1]试求使效用函数q最大的工程计划.

解： 约束方程$4x^2+9y^2=36$描述的不是一个单位向量的集合，但是通过一个变量替换可以做到这一点. 将这个约束重写成：

$$\left(\frac{x}{3}\right)^2+\left(\frac{y}{2}\right)^2=1$$

1. 无差别曲线的讨论详见 Michael D. Intriligator, Ronald G. Bodkin, and Cheng Hsiao, *Econometric Models, Techniques, and Applications* (Upper Saddle River, NJ: Prentice-Hall, 1996).



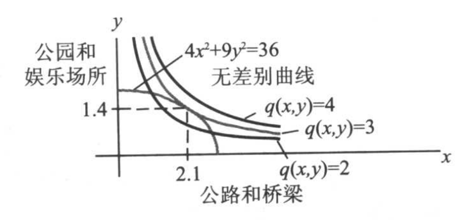


图7-11 最佳的公共工程实施计划是(2.1,1.4)

并且定义

$$x_1=\frac{x}{3},\quad x_2=\frac{y}{2},\quad 即,\quad x=3x_1\quad 且\quad y=2x_2$$

则约束方程变为

$$x_1^2+x_2^2=1$$

效用函数变为 $q(3x_1,2x_2)=(3x_1)(2x_2)=6x_1x_2$. 令 $\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\end{bmatrix}$. 则问题转化为求 $\boldsymbol{Q}(\boldsymbol{x})=6x_1x_2$ 在约束条件 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 下的最大值. 注意 $\boldsymbol{Q}(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$,其中

$$\boldsymbol{A}=\begin{bmatrix}0&3\\3&0\end{bmatrix}$$

469 $\boldsymbol{A}$ 的本征值是 ±3, 且对于 $\lambda=3$ 与 $\lambda=-3$ 分别有本征向量 $\begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$和$\begin{bmatrix}-1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$. 因此 $\boldsymbol{Q}(\boldsymbol{x})=q(x_1,x_2)$的最大值是3, 并且在 $x_1=1/\sqrt{2},x_2=1/\sqrt{2}$时可以取到这个值.

采用原来变量的记法, 最优的公共工程实施计划是 $x=3x_1=3/\sqrt{2}\approx2.1$ 百里公路和桥梁, 以及 $y=2x_2=\sqrt{2}\approx1.4$ 百亩公园和休闲娱乐场所. 这个最优计划恰好是约束曲线与无差别曲线 $q(x,y)=3$ 的交点. 有着更高效用的点(x,y)落在与约束曲线不相交的那些无差别曲线上. 见图7-11. ■

基础练习

1. 设 $\boldsymbol{Q}(\boldsymbol{x})=3x_1^2+3x_2^2+2x_1x_2$. 求一个变量替换, 它将 $\boldsymbol{Q}$ 变换成不含交叉乘积项的二次型, 并给出新的二次型.
2. 设 $\boldsymbol{Q}$ 同基础练习1, 求 $\boldsymbol{Q}(\boldsymbol{x})$在约束条件 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 下的最大值, 以及使 $\boldsymbol{Q}(\boldsymbol{x})$取到最大值的一个单位向量.

习题 7.3

求习题1和2中将二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 变为 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{y}$ 的变量替换:

1. $5x_1^2+6x_2^2+7x_3^2+4x_1x_2-4x_2x_3=9y_1^2+6y_2^2+3y_3^2$
2. $3x_1^2+2x_2^2+2x_3^2+2x_1x_2+2x_1x_3+4x_2x_3=5y_1^2+2y_2^2$

   [提示: $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的坐标分量个数必定相等, 因此这里所给二次型的 y_3^2 项系数为零.]

习题 3 ~6 中，试求(a)$Q(x)$在约束条件 $x^Tx=1$ 下的最大值；(b)使 $Q(x)$取到最大值的一个单位向量 u；以及(c)$Q(x)$在约束条件 $x^Tx=1$ 和 $x^Tu=1$ 下的最大值.

3. $Q(x)=5x_1^2+6x_2^2+7x_3^2+4x_1x_2-4x_2x_3$(见习题 1.)
4. $Q(x)=3x_1^2+2x_2^2+2x_3^2+2x_1x_2+2x_1x_3+4x_2x_3$(见习题 2.)
5. $Q(x)=5x_1^2+5x_2^2-4x_1x_2$
6. $Q(x)=7x_1^2+3x_2^2+3x_1x_2$
7. 设 $Q(x)=-2x_1^2-x_2^2+4x_1x_2+4x_2x_3$. 求约束条件 $x^Tx=1$ 下，使 $Q(x)$取最大值的一个 $\mathbf{R}^n$ 中的单位向量. [提示：二次型 Q 对应矩阵的本征值是 2，-1 和 -4.]
8. 设 $Q(x)=7x_1^2+x_2^2+7x_3^2-8x_1x_2-4x_1x_3-8x_2x_3$. 求约束条件 $x^Tx=1$ 下，使 $Q(x)$取最大值的一个 $\mathbf{R}^n$ 中的单位向量. [提示：二次型 Q 对应矩阵的本征值是 9 和 -3.]
9. 求 $Q(x)=7x_1^2+3x_2^2-2x_1x_2$ 在约束条件 $x_1^2+x_2^2=1$ 下的最大值.（不用求使 $Q(x)$取到最大值的向量.）
10. 求 $Q(x)=-3x_1^2+5x_2^2-2x_1x_2$ 在约束条件 $x_1^2+x_2^2=1$ 下的最大值.（不用求使 $Q(x)$取到最大值的向量.）
11. 设 x 是矩阵 A 的一个单位向量，并且对应于本征值 3. 则 x^TAx 的值是多少?
12. 设 λ 是对称矩阵 A 的任意一个本征值. 验证本节的论断：$m\leqslant\lambda\leqslant M$，其中 m 和 M 定义同式(2). [提示：求满足 $\lambda=x^TAx$ 的 x.]
13. 设 A 是一个 $n\times n$ 对称矩阵，M 和 m 分别表示二次型 x^TAx 的最大值和最小值，并且以 u_1 和 u_n 记对应的单位本征向量. 下面的推导表明，任意给定 M 和 m 之间的数 t，存在一个单位向量 x，使得 $t=x^TAx$. 验证对 0 到 1 之间的某个 α，有 $t=(1-\alpha)m+\alpha M$. 接下来，令 $x=\sqrt{1-\alpha}u_n+\sqrt{\alpha}u_1$，证明 $x^Tx=1$ 及 $x^TAx=t$.

[M]根据习题 3 ~6 的说明，解习题 14 ~17.

470

14. $x_1x_2+3x_1x_3+30x_1x_4+30x_2x_3+3x_2x_4+x_3x_4$
15. $3x_1x_2+5x_1x_3+7x_1x_4+7x_2x_3+5x_2x_4+3x_3x_4$
16. $4x_1^2-6x_1x_2-10x_1x_3-10x_1x_4-6x_2x_3-6x_2x_4-2x_3x_4$
17. $-6x_1^2-10x_2^2-13x_3^2-13x_4^2-4x_1x_2-4x_1x_3-4x_1x_4+6x_3x_4$

基础练习答案

1. 二次型的矩阵是 $A=\begin{bmatrix}3&1\\1&3\end{bmatrix}$. 易求得其本征值 4 和 2，以及对应的单位本征向量 $\begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$和$\begin{bmatrix}-1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$. 所以所求变量替换为 $x=Py$，其中 $P=\begin{bmatrix}1/\sqrt{2}&-1\sqrt{2}\\1/\sqrt{2}&1/\sqrt{2}\end{bmatrix}$.（这里一个常见的错误是没有标准化本征向量.）新的二次型是 $y^TDy=4y_1^2+2y_2^2$.

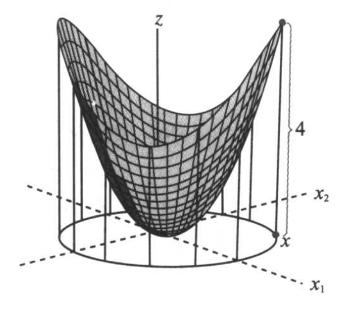


图 7-12 $Q(x)$在约束条件 $x^Tx=1$ 下的最大值为 4

2. $Q(x)$对于单位向量 x 的最大值是 4（如图 7-12），并且这个值在单位本征向量 $\begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$处被取到. [一

个常见的错误答案是$\begin{bmatrix}1\\0\end{bmatrix}$. 这个向量最大化二次型 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{y}$，而不是 $\boldsymbol{Q}(\boldsymbol{x})$.]

7.4　奇异值分解

5.3 和 7.1 节中的对角化定理在许多有趣的实例中都有应用. 遗憾的是，并非所有的矩阵都可以分解成 $\boldsymbol{A}=\boldsymbol{PDP}^{-1}$，其中 $\boldsymbol{D}$ 是对角阵. 不过，任意 $m\times n$ 矩阵 $\boldsymbol{A}$ 都有 $\boldsymbol{A}=\boldsymbol{QDP}^{-1}$形式的分解！这类特殊的分解被称作奇异值分解，它是线性代数中应用最广泛的矩阵分解之一.

奇异值分解基于一般对角化的下述性质：对称矩阵本征值的绝对值度量了将向量(本征向量)拉伸或者收缩的倍数. 并且这种对角化还可以在非方阵的矩形矩阵上加以效仿. 如果 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$ 且 $\|\boldsymbol{x}\|=1$，则

$$\|\boldsymbol{Ax}\|=\|\lambda\boldsymbol{x}\|=|\lambda|\,\|\boldsymbol{x}\|=|\lambda| \tag{1}$$

如果 λ_1是绝对值最大的本征值，则 $\boldsymbol{A}$ 在对应单位本征向量 $\boldsymbol{v}_1$ 方向上的拉伸效果最明显. 也就是说，$\boldsymbol{x}=\boldsymbol{v}_1$时 $\boldsymbol{Ax}$ 的长度取到最大值，且根据(1)有$\|\boldsymbol{Av}_1\|=|\lambda_1|$. 类推到矩形矩阵上，对 $\boldsymbol{v}_1$以及$|\lambda_1|$的描述将导出奇异值分解.

【例题 1】　若 $\boldsymbol{A}=\begin{bmatrix}4&11&14\\8&7&-2\end{bmatrix}$，则线性变换 $\boldsymbol{x}\mapsto\boldsymbol{Ax}$ 将 $\mathbf{R}^3$的单位球面$\{\boldsymbol{x}:\|\boldsymbol{x}\|=1\}$映到 $\mathbf{R}^2$ 中的一个椭圆，如图 7-13. 求最大化长度$\|\boldsymbol{Ax}\|$的一个单位向量，并且算出这个最大长度.

471

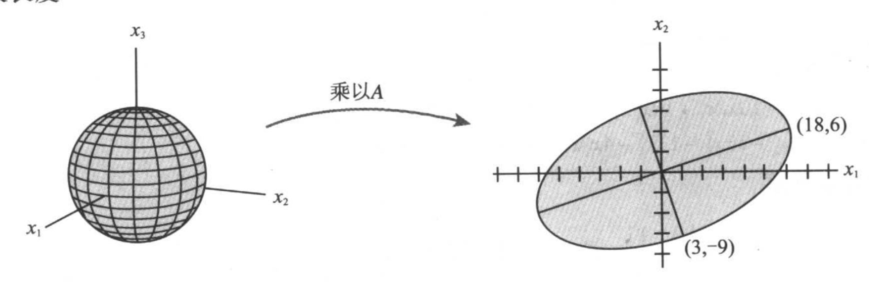


图 7-13　一个从 $\mathbf{R}^3$ 到 $\mathbf{R}^2$ 的变换

解： $\|\boldsymbol{Ax}\|^2$ 与$\|\boldsymbol{Ax}\|$在相同的 $\boldsymbol{x}$ 处取到最大值，而$\|\boldsymbol{Ax}\|^2$ 更容易讨论. 观察到

$$\|\boldsymbol{Ax}\|^2=(\boldsymbol{Ax})^{\mathrm{T}}(\boldsymbol{Ax})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{x}$$

且由$(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{TT}}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$，知 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是对称矩阵. 所以现在的问题是在约束条件$\|\boldsymbol{x}\|=1$下，最大化二次型 $\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{x}$. 根据定理6，这个最大值是$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的最大本征值 λ_1，且这个值在 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的对应于 λ_1的单位本征向量处被取到.

对本题中的矩阵 $\boldsymbol{A}$，有

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{bmatrix}4&8\\11&7\\14&-2\end{bmatrix}\begin{bmatrix}4&11&14\\8&7&-2\end{bmatrix}=\begin{bmatrix}80&100&40\\100&170&140\\40&140&200\end{bmatrix}$$

$\boldsymbol{A}$ 的本征值是 $\lambda_1=360,\lambda_2=90$ 及 $\lambda_3=0$. 对应的单位本征向量分别是：

$$v_1=\begin{bmatrix}1/3\\2/3\\2/3\end{bmatrix},\quad v_2=\begin{bmatrix}-2/3\\-1/3\\2/3\end{bmatrix},\quad v_3=\begin{bmatrix}2/3\\-2/3\\1/3\end{bmatrix}$$

于是$\|Ax\|^2$的最大值为360，并且当x是单位向量v_1时，可以取到这个值．向量Av_1是图7-13中椭圆上与原点相距最远的点，即

$$Av_1=\begin{bmatrix}4&11&14\\8&7&-2\end{bmatrix}\begin{bmatrix}1/3\\2/3\\2/3\end{bmatrix}=\begin{bmatrix}18\\6\end{bmatrix}$$

所以在约束条件$\|x\|=1$下，$\|Ax\|$的最大值是$\|Av_1\|=\sqrt{360}=6\sqrt{10}$. ■

例题1表明A在$\mathbf{R}^3$中单位球面上的影响与二次型$x^{\mathrm{T}}(A^{\mathrm{T}}A)x$有关．实际上，正如我们将看到的，线性变换$x\mapsto Ax$的全部几何性质都由这个二次型所决定．

472

7.4.1 $m\times n$矩阵的奇异值分解

设A是一个$m\times n$矩阵．则$A^{\mathrm{T}}A$对称并且可正交对角化．设$\{v_1,\cdots,v_n\}$是由$A^{\mathrm{T}}A$的本征向量构成的$\mathbf{R}^n$的一组标准正交基，$\lambda_1,\cdots,\lambda_n$是其对应本征值．则对$1\leqslant i\leqslant n$，有

$$\begin{aligned}\|Av_i\|^2&=(Av_i)^{\mathrm{T}}Av_i=v_i^{\mathrm{T}}A^{\mathrm{T}}Av_i\\&=v_i^{\mathrm{T}}(\lambda_i v_i)\quad\text{因为 } v_i \text{ 是 } A^{\mathrm{T}}A \text{ 的本征向量}\\&=\lambda_i\quad\text{因为 } v_i \text{ 是单位向量}\end{aligned}\tag{2}$$

所以$A^{\mathrm{T}}A$的本征值都是非负的．必要时对本征值重新编号，我们可以假定它们排列如下：

$$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n\geqslant0$$

A的**奇异值**(singular values)就是$A^{\mathrm{T}}A$本征值的平方根，记作$\sigma_1,\cdots,\sigma_n$，它们按递降排列．即，对$1\leqslant i\leqslant n$有$\sigma_i=\sqrt{\lambda_i}$．由(2)可知，$A$的奇异值就是向量$Av_1,\cdots,Av_n$的长度．

【例题2】 设A是例题1中的矩阵．由于$A^{\mathrm{T}}A$的本征值是360,90和0,A的奇异值是

$$\sigma_1=\sqrt{360}=6\sqrt{10},\quad\sigma_2=\sqrt{90}=3\sqrt{10},\quad\sigma_3=0$$

由例题1知，A的第一个奇异值是$\|Ax\|$在全体单位向量上的最大值，并且这个值在单位本征向量v_1处被取到．7.3节定理7表明，A的第二个奇异值是$\|Ax\|$在全体正交于v_1的单位向量上的最大值，并且这个值在第二个单位本征向量v_2处被取到(习题22)．对于例题1中的v_2，

$$Av_2=\begin{bmatrix}4&11&14\\8&7&-2\end{bmatrix}\begin{bmatrix}-2/3\\-1/3\\2/3\end{bmatrix}=\begin{bmatrix}3\\-9\end{bmatrix}$$

这个点落在图7-13中椭圆的短轴上，而Av_1则落在长轴上(见图7-14)．A的前两个奇异值分别是椭圆长、短半轴的长度． ■

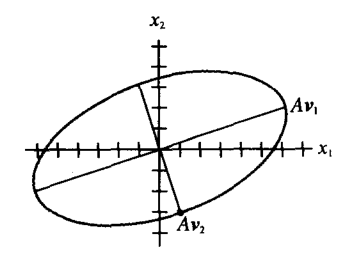


图 7-14

图7-14中Av_1与Av_2正交，这一事实并非偶然，下面定理将予以说明．

【定理9】　设$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n\}$是由$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$的本征向量所构成的$\mathbf{R}^n$的一组标准正交基，其排列使得$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$的对应本征值满足$\lambda_1 \geqslant \cdots \geqslant \lambda_n$，设$\boldsymbol{A}$有$r$个非零奇异值. 则$\{\boldsymbol{A}\boldsymbol{v}_1,\cdots,\boldsymbol{A}\boldsymbol{v}_r\}$是Col $\boldsymbol{A}$的一组正交基，且rank $\boldsymbol{A}=r$.

473

证明：由于$i \neq j$时$\boldsymbol{v}_i$与$\lambda_j\boldsymbol{v}_j$正交，所以

$$(\boldsymbol{A}\boldsymbol{v}_i)^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{v}_j)=\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_j=\boldsymbol{v}_i^{\mathrm{T}}(\lambda_j\boldsymbol{v}_j)=0$$

于是$\{\boldsymbol{A}\boldsymbol{v}_1,\cdots,\boldsymbol{A}\boldsymbol{v}_n\}$是一个正交集. 更进一步，由于向量$\boldsymbol{A}\boldsymbol{v}_1,\cdots,\boldsymbol{A}\boldsymbol{v}_n$的长度是$\boldsymbol{A}$的奇异值，并且存在$r$个非零奇异值，所以$\boldsymbol{A}\boldsymbol{v}_i \neq \boldsymbol{0}$当且仅当$1 \leqslant i \leqslant r$. 因此$\boldsymbol{A}\boldsymbol{v}_1,\cdots,\boldsymbol{A}\boldsymbol{v}_r$线性无关，并且都属于Col $\boldsymbol{A}$. 最后，对于Col $\boldsymbol{A}$中的任意$\boldsymbol{y}$，比如说$\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}$，我们可以记$\boldsymbol{x}=c_1\boldsymbol{v}_1+\cdots+c_n\boldsymbol{v}_n$，以及

$$\begin{aligned}\boldsymbol{y}&=\boldsymbol{A}\boldsymbol{x}=c_1\boldsymbol{A}\boldsymbol{v}_1+\cdots+c_r\boldsymbol{A}\boldsymbol{v}_r+c_{r+1}\boldsymbol{A}\boldsymbol{v}_{r+1}+\cdots+c_n\boldsymbol{A}\boldsymbol{v}_n\\&=c_1\boldsymbol{A}\boldsymbol{v}_1+\cdots+c_r\boldsymbol{A}\boldsymbol{v}_r+0+\cdots+0\end{aligned}$$

于是$\boldsymbol{y}$属于Span$\{\boldsymbol{A}\boldsymbol{v}_1,\cdots,\boldsymbol{A}\boldsymbol{v}_r\}$，这就表明$\{\boldsymbol{A}\boldsymbol{v}_1,\cdots,\boldsymbol{A}\boldsymbol{v}_r\}$是Col $\boldsymbol{A}$的一组（正交）基. 所以rank $\boldsymbol{A}=\dim$ Col $\boldsymbol{A}=r$. ■

注记

某些情况下，$\boldsymbol{A}$的秩对其中元素的微小变化很敏感. 使用计算机对$\boldsymbol{A}$进行行化简时，计数$\boldsymbol{A}$主元列的那种显而易见的方法有时不奏效，因为舍入误差往往会导出一个满秩的阶梯形式.

实际应用中，估计大型矩阵$\boldsymbol{A}$秩的最可靠的方法是数非零奇异值的个数. 此时，出于应用的目的，非常小的非零奇异值将被看成零，矩阵的有效秩是其余非零奇异值的个数.[1]

7.4.2　奇异值分解

$\boldsymbol{A}$的分解包含一个如下形式的$m \times n$"对角"矩阵：

$$\Sigma=\begin{bmatrix}\boldsymbol{D} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix} \begin{matrix}\\ \leftarrow m-r\text{行}\end{matrix} \qquad \uparrow n-r\text{列} \tag{3}$$

其中$\boldsymbol{D}$是$r \times r$对角阵，r不超过m和n中较小者.（如果r等于m或n，或者同时等于m和n，则(3)中部分或者全部零矩阵可能不出现.）

474

【定理10】　奇异值分解

设$\boldsymbol{A}$是秩为r的$m \times n$矩阵. 则存在如(3)所示的$m \times n$矩阵Σ，其中$\boldsymbol{D}$的对角线元素是$\boldsymbol{A}$的前r个奇异值：$\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > 0$，并且存在一个$m \times m$正交矩阵$\boldsymbol{U}$，以及一个$n \times n$正交矩阵$\boldsymbol{V}$，使得

$$\boldsymbol{A}=\boldsymbol{U}\Sigma\boldsymbol{V}^{\mathrm{T}}$$

如果分解式$\boldsymbol{A}=\boldsymbol{U}\Sigma\boldsymbol{V}^{\mathrm{T}}$满足$\boldsymbol{U}$和$\boldsymbol{V}$是正交矩阵，$\Sigma$同(3)中且$\boldsymbol{D}$的对角线元素为

---

1. 通常，秩的估计不是一个简单的问题，有关它的详细讨论见 Phlip E. Gill, Walter Murray, and Margaret H. Wright, *Numerical Linear Algebra and Optimization*, vol. 1 (Redwood City, CA: Addison-Wesley, 1991), Sec. 5.8.

正，则称该分解为 A 的**奇异值分解**(singular value decomposition,或称 SVD)．矩阵 U 和 V 不由 A 唯一确定，但 Σ 的对角线元素一定是 A 的奇异值．见习题 19．称该分解中 U 的列为 A 的**左奇异向量**(left singular vectors)，V 的列为 A 的**右奇异向量**(right singular vectors)．

证明： 设 λ_i 和 $\boldsymbol{v}_i$ 同定理 9，于是 $\{A\boldsymbol{v}_1,\cdots,A\boldsymbol{v}_r\}$ 是 Col A 的一组正交基．标准化每个 $A\boldsymbol{v}_i$ 可以得到标准正交基 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_r\}$，其中

$$\boldsymbol{u}_i=\frac{1}{\|A\boldsymbol{v}_i\|}A\boldsymbol{v}_i=\frac{1}{\sigma_i}A\boldsymbol{v}_i$$

且

$$A\boldsymbol{v}_i=\sigma_i\boldsymbol{u}_i\quad(1\leqslant i\leqslant r)\tag{4}$$

现在将 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_r\}$ 扩充成 $\mathbf{R}^m$ 的一组标准正交基 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_m\}$，并且令

$$U=[\boldsymbol{u}_1\quad \boldsymbol{u}_2\quad \cdots\quad \boldsymbol{u}_m]\quad 以及\ V=[\boldsymbol{v}_1\quad \boldsymbol{v}_2\quad \cdots\quad \boldsymbol{v}_n]$$

由上述构造可知，U 和 V 是正交矩阵．此外，由(4)有

$$AV=[A\boldsymbol{v}_1\quad \cdots\quad A\boldsymbol{v}_r\quad \boldsymbol{0}\quad \cdots\quad \boldsymbol{0}]=[\sigma_1\boldsymbol{u}_1\quad \cdots\quad \sigma_r\boldsymbol{u}_r\quad \boldsymbol{0}\quad \cdots\quad \boldsymbol{0}]$$

令 D 是对角线元素为 $\sigma_1,\cdots,\sigma_r$ 的对角阵，令 Σ 同前文(3)．则

$$U\Sigma=[\boldsymbol{u}_1\quad \boldsymbol{u}_2\quad \cdots\quad \boldsymbol{u}_m]\left[\begin{array}{cccc|c}\sigma_1 & & & \boldsymbol{0} & \\ & \sigma_2 & & & \boldsymbol{0}\\ & & \ddots & & \\ \boldsymbol{0} & & & \sigma_r & \\ \hline & & \boldsymbol{0} & & \boldsymbol{0}\end{array}\right]=[\sigma_1\boldsymbol{u}_1\quad \cdots\quad \sigma_r\boldsymbol{u}_r\quad \boldsymbol{0}\quad \cdots\quad \boldsymbol{0}]=AV$$

由于 V 是正交矩阵，所以 $U\Sigma V^{\mathrm{T}}=AVV^{\mathrm{T}}=A$． ■ [475]

下面两个例题将注意力集中到奇异值分解的内部结构上．还有一类高效且数值稳定的分解算法采用了另一种不同途径．见本节末尾的注记．

【例题 3】 利用例题 1 和 2 的结果，构造 $A=\begin{bmatrix}4 & 11 & 14\\ 8 & 7 & -2\end{bmatrix}$ 的一个奇异值分解．

解： 构造过程可以分成三个步骤：

步骤 1： 求 $A^{\mathrm{T}}A$ 的一个正交对角化形式．即求 $A^{\mathrm{T}}A$ 的本征值，及其对应本征向量所构成的一个标准正交集．如果 A 只包含两列，则可以用手算来完成计算．对于更大的矩阵，通常需要利用矩阵软件．[1]不过对于这里的矩阵 A，例题 1 已经提供了的本征数据．

步骤 2： 建立 V 和 Σ．将 $A^{\mathrm{T}}A$ 的本征值按降序排列．由例题 1 的解答可知，它们分别为：360，90 和 0．其对应的单位本征向量 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 和 $\boldsymbol{v}_3$ 就是 A 的右奇异向量．利用例题 1，构造

$$V=[\boldsymbol{v}_1\quad \boldsymbol{v}_2\quad \boldsymbol{v}_3]=\begin{bmatrix}1/3 & -2/3 & 2/3\\ 2/3 & -1/3 & -2/3\\ 2/3 & 2/3 & 1/3\end{bmatrix}$$

本征值的平方根就是奇异值：

$$\sigma_1=6\sqrt{10},\quad \sigma_2=3\sqrt{10},\quad \sigma_3=0$$

---

1. 关于矩阵软件和图形计算器的命令，可以参考学习指南．例如，MATLAB 中单个命令 eig 可以同时求出本征值和本征向量．

非零奇异值是 $\boldsymbol{D}$ 的对角元素. 矩阵 Σ 的维度与 $\boldsymbol{A}$ 相同, $\boldsymbol{D}$ 位于其左上角, 其余元素皆为0.

$$\boldsymbol{D}=\begin{bmatrix}6\sqrt{10} & 0\\ 0 & 3\sqrt{10}\end{bmatrix},\quad \Sigma=[\boldsymbol{D}\quad \boldsymbol{0}]=\begin{bmatrix}6\sqrt{10} & 0 & 0\\ 0 & 3\sqrt{10} & 0\end{bmatrix}$$

步骤3: 构造 $\boldsymbol{U}$. 当 $\boldsymbol{A}$ 的秩为 r 时, $\boldsymbol{U}$ 的前 r 列是 $\boldsymbol{A}\boldsymbol{v}_1,\cdots,\boldsymbol{A}\boldsymbol{v}_r$ 经标准化所得的向量. 本例中, $\boldsymbol{A}$ 有两个非零奇异值, 所以 $\operatorname{rank}\boldsymbol{A}=2$. 回忆(2)式以及例题2前面的讨论, 有 $\|\boldsymbol{A}\boldsymbol{v}_1\|=\sigma_1$, $\|\boldsymbol{A}\boldsymbol{v}_2\|=\sigma_2$. 于是

$$\boldsymbol{u}_1=\frac{1}{\sigma_1}\boldsymbol{A}\boldsymbol{v}_1=\frac{1}{6\sqrt{10}}\begin{bmatrix}18\\ 6\end{bmatrix}=\begin{bmatrix}3/\sqrt{10}\\ 1/\sqrt{10}\end{bmatrix}$$

$$\boldsymbol{u}_2=\frac{1}{\sigma_2}\boldsymbol{A}\boldsymbol{v}_2=\frac{1}{3\sqrt{10}}\begin{bmatrix}3\\ -9\end{bmatrix}=\begin{bmatrix}1/\sqrt{10}\\ -3/\sqrt{10}\end{bmatrix}$$

注意到 $\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$ 已经是 $\mathbf{R}^2$ 的一组基. 所以为了构造 $\boldsymbol{U}$ 不需要其他向量, 令 $\boldsymbol{U}=[\boldsymbol{u}_1\quad \boldsymbol{u}_2]$ 即可. $\boldsymbol{A}$ 的奇异值分解是

476

$$\boldsymbol{A}=\underset{\uparrow\atop \boldsymbol{U}}{\begin{bmatrix}3/\sqrt{10} & 1/\sqrt{10}\\ 1/\sqrt{10} & -3/\sqrt{10}\end{bmatrix}}\underset{\uparrow\atop \Sigma}{\begin{bmatrix}6\sqrt{10} & 0 & 0\\ 0 & 3\sqrt{10} & 0\end{bmatrix}}\underset{\uparrow\atop \boldsymbol{V}^{\mathrm{T}}}{\begin{bmatrix}1/3 & 2/3 & 2/3\\ -2/3 & -1/3 & 2/3\\ 2/3 & -2/3 & 1/3\end{bmatrix}}$$

■

【**例题4**】　求 $\boldsymbol{A}=\begin{bmatrix}1 & -1\\ -2 & 2\\ 2 & -2\end{bmatrix}$ 的一个奇异值分解.

解: 首先计算 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{bmatrix}9 & -9\\ -9 & 9\end{bmatrix}$. $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的本征值是18和0, 其对应单位本征向量为:

$$\boldsymbol{v}_1=\begin{bmatrix}1/\sqrt{2}\\ -1/\sqrt{2}\end{bmatrix},\quad \boldsymbol{v}_2=\begin{bmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}$$

这些单位向量构成 $\boldsymbol{V}$ 的列:

$$\boldsymbol{V}=[\boldsymbol{v}_1\quad \boldsymbol{v}_2]=\begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\ -1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}$$

奇异值为 $\sigma_1=\sqrt{18}=3\sqrt{2}$, $\sigma_2=0$. 因为只有一个非零奇异值, "矩阵" $\boldsymbol{D}$ 可以写成一个单独的数. 即 $\boldsymbol{D}=3\sqrt{2}$. 矩阵 Σ 的维度与 $\boldsymbol{A}$ 相同, 且 $\boldsymbol{D}$ 位于其左上角:

$$\Sigma=\begin{bmatrix}D & 0\\ 0 & 0\\ 0 & 0\end{bmatrix}=\begin{bmatrix}3\sqrt{2} & 0\\ 0 & 0\\ 0 & 0\end{bmatrix}$$

为构造 $\boldsymbol{U}$，我们先构造 $\boldsymbol{A}\boldsymbol{v}_1$ 和 $\boldsymbol{A}\boldsymbol{v}_2$：

$$\boldsymbol{A}\boldsymbol{v}_1=\begin{bmatrix}2/\sqrt{2}\\-4/\sqrt{2}\\4/\sqrt{2}\end{bmatrix},\quad \boldsymbol{A}\boldsymbol{v}_2=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

为检验前面计算的正确性，验证 $\|\boldsymbol{A}\boldsymbol{v}_1\|=\sigma_1=3\sqrt{2}$. 由 $\|\boldsymbol{A}\boldsymbol{v}_2\|=\sigma_2=0$ 知 $\boldsymbol{A}\boldsymbol{v}_2=\boldsymbol{0}$. 到此，仅得到 $\boldsymbol{U}$ 的一列：

$$\boldsymbol{u}_1=\frac{1}{3\sqrt{2}}\boldsymbol{A}\boldsymbol{v}_1=\begin{bmatrix}1/3\\-2/3\\2/3\end{bmatrix}$$

$\boldsymbol{U}$ 的其余各列可以通过将 $\{\boldsymbol{u}_1\}$ 扩张成 $\mathbf{R}^3$ 的标准正交基来得到. 这时我们需要两个正交的单位向量 $\boldsymbol{u}_2$ 和 $\boldsymbol{u}_3$，并且它们都与 $\boldsymbol{u}_1$ 正交（见图 7-15）. 每个向量都满足 $\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{x}=0$，或者等价地，满足方程 $x_1-2x_2+2x_3=0$. 此方程解集的一组基是：

$$\boldsymbol{w}_1=\begin{bmatrix}2\\1\\0\end{bmatrix},\quad \boldsymbol{w}_2=\begin{bmatrix}-2\\0\\1\end{bmatrix}$$

检查 $\boldsymbol{w}_1$ 和 $\boldsymbol{w}_2$ 都与 $\boldsymbol{u}_1$ 正交. 对 $\{\boldsymbol{w}_1,\ \boldsymbol{w}_2\}$ 运用格拉姆－施密特方法（含标准化），可得

$$\boldsymbol{u}_2=\begin{bmatrix}2/\sqrt{5}\\1/\sqrt{5}\\0\end{bmatrix},\quad \boldsymbol{u}_3=\begin{bmatrix}-2/\sqrt{45}\\4/\sqrt{45}\\5/\sqrt{45}\end{bmatrix}$$

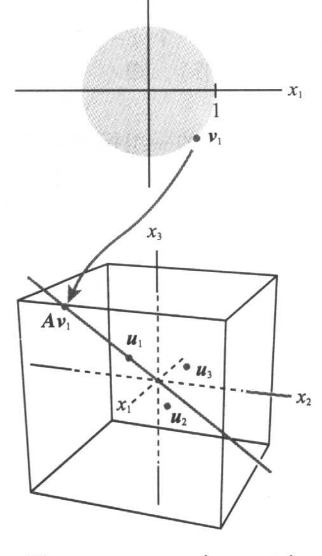


图 7-15 $\boldsymbol{u}_2$、$\boldsymbol{u}_3$ 与 $\boldsymbol{u}_1$ 正交

477

最后，令 $\boldsymbol{U}=[\boldsymbol{u}_1\boldsymbol{u}_2\boldsymbol{u}_3]$，取上面的 Σ 和 $\boldsymbol{V}^{\mathrm{T}}$，记

$$\boldsymbol{A}=\begin{bmatrix}1&-1\\-2&2\\2&-2\end{bmatrix}=\begin{bmatrix}1/3&2/\sqrt{5}&-2/\sqrt{45}\\-2/3&1/\sqrt{5}&4/\sqrt{45}\\2/3&0&5/\sqrt{45}\end{bmatrix}\begin{bmatrix}3/\sqrt{2}&0\\0&0\\0&0\end{bmatrix}\begin{bmatrix}1/\sqrt{2}&-1/\sqrt{2}\\ \\1/\sqrt{2}&1/\sqrt{2}\end{bmatrix}$$ ■

7.4.3 奇异值分解的应用

如前文所述，SVD 常用于估计矩阵的秩. 下面简单介绍它的其他几种数值应用，在图像处理上的应用将在 7.5 节给出.

【例题 5】 （条件数）涉及 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的大多数数值计算，当采用 $\boldsymbol{A}$ 的 SVD 时最为可靠. 两个正交矩阵 $\boldsymbol{U}$ 和 $\boldsymbol{V}$ 不会改变向量的长度以及向量间的夹角（6.2 节定理 7）. 数值计算中所有可能的不稳定因素就只有 Σ. 如果 $\boldsymbol{A}$ 的奇异值非常大或者非常小，舍入误差几乎不可避免，但若知道 Σ 和 $\boldsymbol{V}$ 中的元素将会有助于我们进行误差分析.

如果 $\boldsymbol{A}$ 是一个 $n\times n$ 可逆矩阵，则最大奇异值与最小奇异值的比值 σ_1/σ_n 就是 $\boldsymbol{A}$ 的**条件数**（condition number）. 2.3 节习题 41 ~ 43 向我们展示了，条件数如何影响 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的解对 $\boldsymbol{A}$ 中元素变化（或误差）的敏感程度. （实际上，$\boldsymbol{A}$ 的“条件数”可以用

多种方法来计算，不过这里给出的定义广泛用于 $\boldsymbol{Ax}=\boldsymbol{b}$ 的研究.）■

【例题 6】　(基本子空间的基)给定 $m\times n$ 矩阵 $\boldsymbol{A}$ 的一个 SVD，令 $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_m$ 为左奇异向量，$\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n$ 为右奇异向量，$\sigma_1,\cdots,\sigma_n$ 为奇异值，且 r 为 $\boldsymbol{A}$ 的秩. 根据定理 9，

$$\{\boldsymbol{u}_1,\dots,\boldsymbol{u}_r\} \tag{5}$$

是 Col $\boldsymbol{A}$ 的一组标准正交基.

回忆 6.1 节定理 3 有 $(\text{Col}\,\boldsymbol{A})^{\perp}=\text{Nul}\,\boldsymbol{A}^{\mathrm{T}}$. 所以

$$\{\boldsymbol{u}_{r+1},\dots,\boldsymbol{u}_m\} \tag{6}$$

是 Nul $\boldsymbol{A}^{\mathrm{T}}$ 的一组标准正交基.

由于对 $1\leqslant i\leqslant n$ 有 $\|\boldsymbol{Av}_i\|=\sigma_i$，并且 σ_i 是 0 当且仅当 $i>r$，所以向量 $\boldsymbol{v}_{r+1},\cdots,\boldsymbol{v}_n$ 张成 Nul $\boldsymbol{A}$ 的一个 $n-r$ 维子空间. 根据秩定理有 dim Nul $\boldsymbol{A}=n-\text{rank}\,\boldsymbol{A}$. 于是，根据基定理(4.5 节)，知

478

$$\{\boldsymbol{v}_{r+1},\dots,\boldsymbol{v}_n\} \tag{7}$$

是以 Nul $\boldsymbol{A}$ 的一组标准正交基.

由(5)和(6)可知，Nul $\boldsymbol{A}^{\mathrm{T}}$ 的正交补是 Col $\boldsymbol{A}$. 交换 $\boldsymbol{A}$ 与 $\boldsymbol{A}^{\mathrm{T}}$，我们有 $(\text{Nul}\,\boldsymbol{A})^{\perp}=\text{Col}\,\boldsymbol{A}^{\mathrm{T}}=\text{Row}\,\boldsymbol{A}$. 于是，由(7)知

$$\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_r\} \tag{8}$$

是 Row $\boldsymbol{A}$ 的一组标准正交基.

图 7-17 总结了(5)～(8)，不过其中给出的是 Col $\boldsymbol{A}$ 的正交基 $\{\sigma_1\boldsymbol{u}_1,\ \cdots,\ \sigma_r\boldsymbol{u}_r\}$，而不是标准正交基.，这是为了提醒你对 $1\leqslant i\leqslant r$ 有 $\boldsymbol{Av}_i=\sigma_i\boldsymbol{u}_i$. $\boldsymbol{A}$ 确定的四个基本子空间的显式标准正交基在一些计算、特别是约束优化问题中十分有用. ■

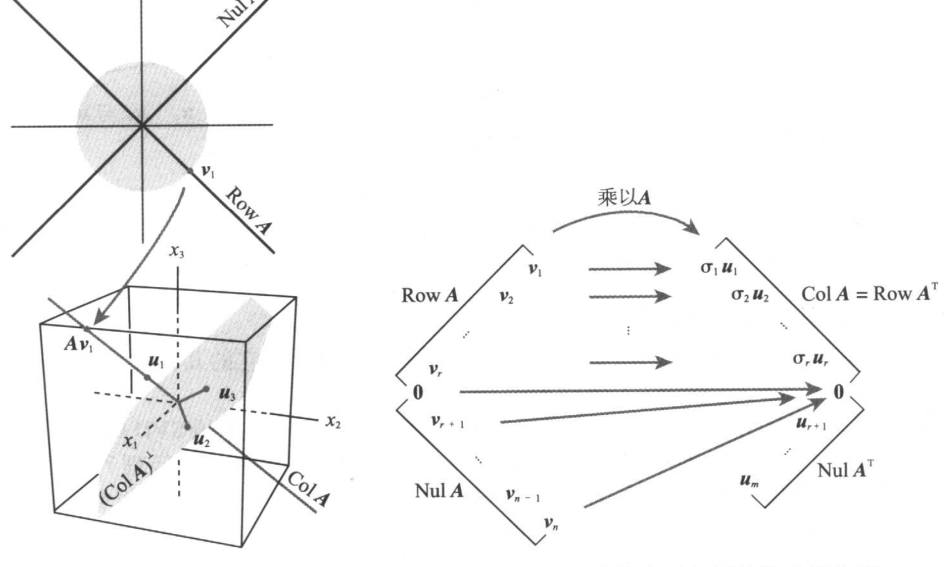


图 7-16　例题 4 中的基本子空间　　　图 7-17　四个基本子空间以及 $\boldsymbol{A}$ 的作用

四个基本子空间以及奇异值的概念给出了可逆矩阵定理的最后陈述. 回忆一下，

为避免命题数目翻倍，A^T的相关命题在此略过. 该定理的其他命题在2.3，2.9，3.2，4.6和5.2节已陆续给出.

【定理11】 可逆矩阵定理(最后结论)

设A是一个$n\times n$矩阵. 则下列各命题都等价于A是可逆矩阵：

u. $(\text{Col } A)^{\perp}=\{\mathbf{0}\}$.　　v. $(\text{Nul } A)^{\perp}=\mathbb{R}^n$.

w. $\text{Row } A=\mathbb{R}^n$.　　x. A有n个非零奇异值

479

【例题7】 (简化SVD和A的伪逆)当Σ包含零行或零列时，A可能有更紧凑的分解. 采用上面的记号，令$r=\text{rank } A$，将U和V划分成子矩阵，其中第1块包含列：

$$U=[U_r \quad U_{m-r}],\text{其中 } U_r=[u_1 \quad \cdots \quad u_r]$$
$$V=[V_r \quad V_{n-r}],\text{其中 } V_r=[v_1 \quad \cdots \quad v_r]$$

则U_r是$m\times r$矩阵，V_r是$n\times r$矩阵.（为简化记号，我们考虑U_{m-r}或V_{n-r}，尽管两者之一有可能不含任何列.）分块矩阵乘法表明，

$$A=[U_r \quad U_{m-r}]\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} V_r^T \\ V_{n-r}^T \end{bmatrix}=U_rDV_r^T \tag{9}$$

A的这种分解称为A的**简化奇异值分解**(reduced singular value decomposition). 由于D的对角线元素非零，我们可以得到下面矩阵，它被称作A的**伪逆**(pseudoinverse)，或**摩尔-彭罗斯逆**(Moore-Penrose inverse)

$$A^{+}=V_rD^{-1}U_r^T \tag{10}$$

本章末尾的补充题12~14探讨了简化奇异值分解及伪逆的一些性质. ■

【例题8】 (最小二乘解)给定方程$Ax=b$，利用(10)中的伪逆，我们定义：

$$\hat{x}=A^{+}b=V_rD^{-1}U_r^Tb$$

则由(9)中的SVD可知

$$\begin{aligned} A\hat{x} &= (U_rDV_r^T)(V_rD^{-1}U_r^Tb) \\ &= U_rDD^{-1}U_r^Tb \quad \text{因为 } V_r^TV_r=I_r \\ &= U_rU_r^Tb \end{aligned}$$

由(5)可知，$U_rU_r^Tb$是b在Col A上的正交投影$\hat{b}$(见6.3节定理10). 于是$\hat{x}$是$Ax=b$的一个最小二乘解. 事实上，在$Ax=b$的所有最小二乘解中，这个$\hat{x}$长度最小. 见补充题14. ■

注记

例题1~4和习题讲解了奇异值的概念以及手算方法. 在实际应用中，由于A中元素的误差在A^TA的元素上被平方放大，因此应当避免计算A^TA. 存在求A的奇异值和奇异向量的快速迭代算法，可精确到小数点后很多位.

480

更多阅读

Horn, Roger A., and Charles R. Johnson, *Matrix Analysis*, vol. 1 (Cambridge: Cambridge University Press, 1985), pp. 414~445.

Long, Cliff, "Visualization of Matrix Singular Value Decomposition." *Mathematics Magazine* **56**(1983), pp. 161~167.

Moler, C. B., and D. Morrison, "Singular Value Analysis of Cryptograms." *Amer. Math. Monthly* **90** (1983), pp. 78 ~ 87.
Strang, Gilbert, *Linear Algebra and Its Applications*, 3rd ed. (San Diego: Harcourt Brace Jovanovich, 1988), pp. 442 ~ 452.
Watkins, David S., *Fundamentals of Matrix Computations* (New York: Wiley, 1991), pp. 390 ~ 398, 409 ~ 421.

基础练习

已知奇异值分解 $A = U\Sigma V^{T}$. 求 A^{T} 的一个 SVD. A 与 A^{T} 的奇异值之间有什么联系?

习题　7.4

求习题 1 ~ 4 中各矩阵的奇异值.

1. $\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$　2. $\begin{bmatrix} -5 & 0 \\ 0 & 0 \end{bmatrix}$　3. $\begin{bmatrix} \sqrt{6} & 1 \\ 0 & \sqrt{6} \end{bmatrix}$　4. $\begin{bmatrix} \sqrt{3} & 2 \\ 0 & \sqrt{3} \end{bmatrix}$

求习题 5 ~ 12 中各矩阵的 SVD. [提示: 习题 11 中, U 可以选取为 $\begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}$.

习题 12 中, U 的一列可以是 $\begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$.]

5. $\begin{bmatrix} -3 & 0 \\ 0 & 0 \end{bmatrix}$　6. $\begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}$　7. $\begin{bmatrix} 2 & -1 \\ 2 & 2 \end{bmatrix}$　8. $\begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$

9. $\begin{bmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{bmatrix}$　10. $\begin{bmatrix} 4 & -2 \\ 2 & -1 \\ 0 & 0 \end{bmatrix}$　11. $\begin{bmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{bmatrix}$　12. $\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}$

13. 求 $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$ 的 SVD. [提示: 讨论 A^{T}.]

14. 求习题 7 中使 $A\boldsymbol{x}$ 长度最大的单位向量 $\boldsymbol{x}$.

15. 假设下面的分解是矩阵 A 的一个 SVD, 其中 U 和 V 的元素保留两位小数.

$$A = \begin{bmatrix} 0.40 & -0.78 & 0.47 \\ 0.37 & -0.33 & -0.87 \\ -0.84 & -0.52 & -0.16 \end{bmatrix} \begin{bmatrix} 7.10 & 0 & 0 \\ 0 & 3.10 & 0 \\ 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} 0.30 & -0.51 & -0.81 \\ 0.76 & 0.64 & -0.12 \\ 0.58 & -0.58 & 0.58 \end{bmatrix}$$

a. A 的秩是多少?

b. 利用 A 的这个分解, 不通过计算, 分别写出 Col A 和 Nul A 的一组基. [提示: 先写出 V 的列.]

481

16. 对下列 3 × 4 矩阵的 SVD 分解重复习题 15.

$$A = \begin{bmatrix} -0.86 & -0.11 & -0.50 \\ 0.31 & 0.68 & -0.67 \\ 0.41 & -0.73 & -0.55 \end{bmatrix} \begin{bmatrix} 12.48 & 0 & 0 & 0 \\ 0 & 6.34 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} 0.66 & -0.03 & -0.35 & 0.66 \\ -0.13 & -0.90 & -0.39 & -0.13 \\ 0.65 & 0.08 & -0.16 & -0.73 \\ -0.34 & 0.42 & -0.84 & -0.08 \end{bmatrix}$$

习题 17 ~ 24 中，A 是一个有奇异值分解 $A = U\Sigma V^{T}$ 的 $m\times n$ 矩阵，其中 U 是 $m\times m$ 正交矩阵，Σ 是 $m\times n$ “对角” 矩阵，其对角线元素有 r 个正值且无负值，V 是 $n\times n$ 正交矩阵. 求解或证明各题的答案.

17. 设 A 是一个可逆方阵. 求 A^{-1} 的一个奇异值分解.
18. 证明：若 A 是方阵，则 $|\det A|$ 是 A 的奇异值的乘积.
19. 证明：U 的列是 $A^{T}A$ 的本征向量，V 的列是 AA^{T} 的本征向量，Σ 的对角线元素是 A 的奇异值. [提示：利用 SVD 来计算 $A^{T}A$ 和 AA^{T}.]
20. 证明：如果 A 是 $n\times n$ 正定矩阵，则正交对角化 $A = PDP^{T}$ 是 A 的一个奇异值分解.
21. 证明：如果 P 是 $m\times m$ 正交矩阵，则 PA 与 A 有相同的奇异值.
22. 证明例题 2 中的断言：矩阵 A 的第二个奇异值是当 x 取遍全体与 v_1 正交的单位向量时 $\|Ax\|$ 的最大值，其中 v_1 是对应于 A 的第一个奇异值的一个右奇异向量. [提示：利用 7.3 节定理 7.]
23. 如果 $U = [u_1 \ \cdots \ u_m]$ 且 $V = [v_1 \ \cdots \ v_n]$，证明：
$$\boxed{A = \sigma_1 u_1 v_1^{T} + \sigma_2 u_2 v_2^{T} + \cdots + \sigma_r u_r v_r^{T}}$$
24. 采用习题 23 的记号，证明：对 $1\leqslant j\leqslant r = \operatorname{rank} A$，有 $A^{T}u_j = \sigma_j v_j$.
25. 设 $T:\mathbf{R}^n\to\mathbf{R}^m$ 是一个线性变换. 说明如何求 $\mathbf{R}^n$ 的基 $\mathcal{B}$ 以及 $\mathbf{R}^m$ 的基 $\mathcal{C}$，使得 T 在 $\mathcal{B}$ 和 $\mathcal{C}$ 下的矩阵是 $m\times n$ “对角” 矩阵.

[M]利用例题 3 和 4 的方法，计算习题 26 和 27 中各矩阵的一个 SVD. 结果矩阵中的元素保留两位小数.

26. $A = \begin{bmatrix} -18 & 13 & -4 & 4 \\ 2 & 19 & -4 & 12 \\ -14 & 11 & -12 & 8 \\ -2 & 21 & 4 & 8 \end{bmatrix}$

27. $A = \begin{bmatrix} 6 & -8 & -4 & 5 & -4 \\ 2 & 7 & -5 & -6 & 4 \\ 0 & -1 & -8 & 2 & 2 \\ -1 & -2 & 4 & 4 & -8 \end{bmatrix}$

28. [M]计算 2.3 节习题 9 中 4×4 矩阵的奇异值以及条件数 σ_1/σ_4.
29. [M]计算 2.3 节习题 10 中 5×5 矩阵的奇异值以及条件数 σ_1/σ_5.

基础练习答案

若 $A = U\Sigma V^{T}$，其中 Σ 是 $m\times n$ 矩阵，则 $A^{T} = (V^{T})^{T}\Sigma^{T}U^{T} = V\Sigma^{T}U^{T}$. 由于 V 和 U 是正交矩阵，且 Σ^{T} 是 $n\times m$ “对角” 矩阵，因此这是 A^{T} 的一个 SVD. 因为 Σ 和 Σ^{T} 有同样的非零对角线元素，所以 A 和 A^{T} 有同样的非零奇异值. [注意：若 A 是 $2\times n$ 矩阵，则 AA^{T} 是 2×2 矩阵，其本征值比 $A^{T}A$ 的本征值更容易计算(手算).]

7.5 图像处理和统计学中的应用

本章实例介绍中的卫星照片给出了多维数据，或称多元数据的一个例子. 多维数据或多元数据是指用特定的方式组织信息，使得在数据集中的每个数据都等同于 $\mathbf{R}^n$ 中的一个点(向量). 本节的主要目标是介绍多元数据的一种分析方法，称作主成分分析. 它将阐述如何应用正交对角化以及奇异值分解. 482

如果数据由某个对象集或个体集上的观测值列表所构成，那么就可以对这些数据使用主成分分析. 例如，考虑生成某种塑胶材料的化学过程. 为监控这一过程，从生成的材料中取出 300 个样本，每个样本都属于一个试验组，试验量有 8 个，包括熔点、密度、抗张强度等等. 针对每个样本的试验报告是 $\mathbf{R}^8$ 中的一个向量，这些向量的集合形成一个 8×300 矩阵，称作**观测矩阵**(matrix of observations).

我们可以粗略地说，这些过程控制数据是8维数据．下面的两个例题描述的是具有几何直观的数据．

【例题1】　已知一个2维数据集由N个大学生的体重和身高给出．以$\boldsymbol{X}_j$记$\mathbf{R}^2$中包含第j个学生体重和身高的**观测向量**(observation vector)．如果w和h分别表示体重和身高，则观测矩阵有如下形式：

$$\begin{bmatrix} w_1 & w_2 & \cdots & w_N \\ h_1 & h_2 & \cdots & h_N \end{bmatrix}$$
$$\begin{matrix} \uparrow & \uparrow & & \uparrow \\ \boldsymbol{X}_1 & \boldsymbol{X}_2 & & \boldsymbol{X}_N \end{matrix}$$

观测向量的集合可以直观地看作一个2维散列图．见图7-18．■

【例题2】　本章实例介绍中所给的内华达州铁路峡谷的前三张照片，由于是在三个独立波长下对同一地区的同步测量所得，因此可以看成是同一幅图像的三个谱成分．三张照片中，每张照片给出了同一地理区域（例如，每张照片左上角的第一个像素对应与地面上的同一位置，范围约为30米×30米）的不同信息．对每个像素，都存在$\mathbf{R}^3$中的一个观测向量，它列出了这个像素在三个谱段下的信号强度．

483

一幅典型的图像是2000×2000像素，因此图像中包含4百万个像素．图像数据形成一个3行、4百万列的矩阵(列的顺序可以是任意的)．这时，数据的“多维”特征是指这3个谱维度，而不是指任何照片都自然具备的两个空间维度．数据可以直观地看作$\mathbf{R}^3$中4百万个点所构成的簇，图7-19给出了一种可能的图像．■

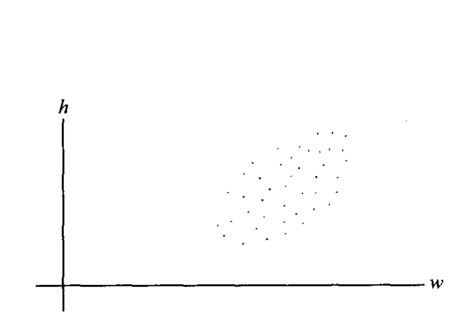


图7-18　观测向量$\boldsymbol{X}_1,\cdots,\boldsymbol{X}_N$所构成的散列图

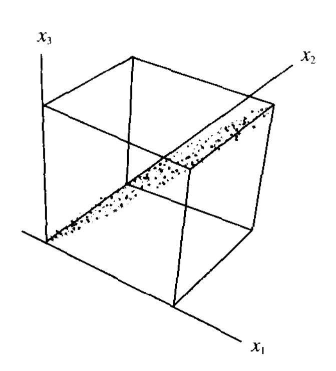


图7-19　卫星图像波谱数据所构成的散列图

7.5.1　平均值与协方差

在进行主成分分析以前，我们先学习几个新概念．设$[\boldsymbol{X}_1 \cdots \boldsymbol{X}_N]$是如前所述的一个$p\times N$观测矩阵．观测向量$\boldsymbol{X}_1,\cdots,\boldsymbol{X}_N$的**样本均值**(sample mean)为

$$\boldsymbol{M}=\frac{1}{N}(\boldsymbol{X}_1+\cdots+\boldsymbol{X}_N)$$

对于图7-18中的数据，样本均值是散列图中“居中”的点．对$k=1,\cdots,N$，令

$$\hat{\boldsymbol{X}}_k=\boldsymbol{X}_k-\boldsymbol{M}$$

$p \times N$ 矩阵

$$B = [\hat{X}_1 \quad \hat{X}_2 \quad \cdots \quad \hat{X}_N]$$

各列的样本均值为零，称 B 具有**平均偏差形式**(mean-deviation form). 将图7-18 中的数据减去其样本均值，所得的散列图如图 7-20.

(样本)协方差矩阵[(sample) covariance matrix]是 $p \times p$ 矩阵 S，其定义为：

$$S = \frac{1}{N-1} BB^{\mathrm{T}}$$

由于任意 BB^{T}形式的矩阵都是半正定的，所以 S 也半正定(见 7.2 节习题 25，交换 B 与 B^{T}的位置).

图 7-20 具有平均偏差形式的体重－身高数据

【例题 3】 已知某种群的一个大小为 4 的随机样本集合，对每一个样本进行三种观测. 观测向量为：

$$X_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 4 \\ 2 \\ 13 \end{bmatrix}, \quad X_3 = \begin{bmatrix} 7 \\ 8 \\ 1 \end{bmatrix}, \quad X_4 = \begin{bmatrix} 8 \\ 4 \\ 5 \end{bmatrix}$$

计算样本均值及协方差矩阵.

解：样本均值为：

484

$$M = \frac{1}{4}\left(\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 4 \\ 2 \\ 13 \end{bmatrix} + \begin{bmatrix} 7 \\ 8 \\ 1 \end{bmatrix} + \begin{bmatrix} 8 \\ 4 \\ 5 \end{bmatrix} \right) = \frac{1}{4}\begin{bmatrix} 20 \\ 16 \\ 20 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \\ 5 \end{bmatrix}$$

从 $X_1, \cdots, X_4$ 中减去样本均值，得到

$$\hat{X}_1 = \begin{bmatrix} -4 \\ -2 \\ -4 \end{bmatrix}, \quad \hat{X}_2 = \begin{bmatrix} -1 \\ -2 \\ 8 \end{bmatrix}, \quad \hat{X}_3 = \begin{bmatrix} 2 \\ 4 \\ -4 \end{bmatrix}, \quad \hat{X}_4 = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$$

以及

$$B = \begin{bmatrix} -4 & -1 & 2 & 3 \\ -2 & -2 & 4 & 0 \\ -4 & 8 & -4 & 0 \end{bmatrix}$$

所以，样本协方差矩阵是：

$$S = \frac{1}{3}\begin{bmatrix} -4 & -1 & 2 & 3 \\ -2 & -2 & 4 & 0 \\ -4 & 8 & -4 & 0 \end{bmatrix}\begin{bmatrix} -4 & -2 & -4 \\ -1 & -2 & 8 \\ 2 & 4 & -4 \\ 3 & 0 & 0 \end{bmatrix}$$

$$= \frac{1}{3}\begin{bmatrix} 30 & 18 & 0 \\ 18 & 24 & -24 \\ 0 & -24 & 96 \end{bmatrix} = \begin{bmatrix} 10 & 6 & 0 \\ 6 & 8 & -8 \\ 0 & -8 & 32 \end{bmatrix}$$ ■

为讨论 $S = [s_{ij}]$中的元素，以 X 记取值在观测向量集上的向量，以 $x_1, \cdots, x_p$记 X 的坐标. 则以数量 x_1为例，其取值范围是 $X_1, \cdots, X_N$的第一个坐标分量所构成的集合.

对于 $j=1,\cdots,p$，S 的对角线元素 s_{jj} 称为 x_j 的**方差**(variance).

x_j 的方差度量了 x_j 取值的分散程度(见习题13). 例题3中，x_1 的方差是10，x_3 的方差是32. 32比10大，这就表明结果向量中第三个元素的取值比第一个元素的取值更分散.

数据的**全变差**(total variance)是矩阵 S 对角线上全体方差之和. 一般地，方阵的对角线元素之和称作矩阵的**迹**(trace)，记作 $\mathrm{tr}(S)$. 于是

$$\{\text{全变差}\}=\mathrm{tr}(S)$$

$i\neq j$ 时，S 中的元素 s_{ij} 称作 x_i 与 x_j 的**协方差**(covariance). 观察例题3，由于 S 的 $(1,3)$-元是0，所以 x_1 与 x_3 间的协方差是0. 统计学家称此时 x_1 与 x_3 **不相关**(uncorrelated). 当变量 $x_1,\cdots,x_p$ 中的大多数或者全体都不相关，即 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_N$ 的协方差矩阵是对角阵或近似对角阵时，对多元数据 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_N$ 的分析可大为简化.

7.5.2 主成分分析

简单起见，假定矩阵 $[\boldsymbol{X}_1\ \cdots\ \boldsymbol{X}_N]$ 已经具有平均偏差形式. 主成分分析的目的是求一个 $p\times p$ 正交矩阵 $P=[\boldsymbol{u}_1\ \cdots\ \boldsymbol{u}_p]$，它所确定的变量替换 $\boldsymbol{X}=P\boldsymbol{Y}$，或

485

$$\begin{bmatrix}x_1\\x_2\\\vdots\\x_p\end{bmatrix}=[\boldsymbol{u}_1\quad \boldsymbol{u}_2\quad\cdots\quad \boldsymbol{u}_p]\begin{bmatrix}y_1\\y_2\\\vdots\\y_p\end{bmatrix}$$

具有下列性质：新变量 $y_1,\cdots,y_p$ 不相关，并且按方差降序排列.

正交的变量替换 $\boldsymbol{X}=P\boldsymbol{Y}$ 意味着每个观测向量 $\boldsymbol{X}_k$ 都有一个"新名字" $\boldsymbol{Y}_k$，使得 $\boldsymbol{X}_k=P\boldsymbol{Y}_k$. 注意 $\boldsymbol{Y}_k$ 是 $\boldsymbol{X}_k$ 关于 P 的列的坐标向量，并且对 $k=1,\cdots,N$，有 $\boldsymbol{Y}_k=P^{-1}\boldsymbol{X}_k=P^{\mathrm{T}}\boldsymbol{X}_k$.

不难验证，对任意正交矩阵 P，$\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_N$ 的协方差矩阵都是 $P^{\mathrm{T}}SP$(习题11). 因此，所求正交矩阵 P 使 $P^{\mathrm{T}}SP$ 为对角阵. 设 D 是一个对角阵，其对角线元素为 S 的本征值 $\lambda_1,\cdots,\lambda_p$，且满足 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_p\geqslant 0$. 另设 P 是以 S 的对应单位本征向量 $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p$ 为列的正交矩阵. 则 $S=PDP^{\mathrm{T}}$，$P^{\mathrm{T}}SP=D$.

协方差矩阵 S 的单位本征向量 $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p$ 称为(观测矩阵中)数据的**主成分**(principal components). **第一主成分**(first principal component)是 S 的最大本征值所对应的本征向量，**第二主成分**(second principal component)是 S 的第二大本征值所对应的本征向量，依此类推.

第一主成分 $\boldsymbol{u}_1$ 按如下方式确定新变量 y_1. 设 $c_1,\cdots,c_p$ 是 $\boldsymbol{u}_1$ 中的元素. 因为 $\boldsymbol{u}_1^{\mathrm{T}}$ 是 P^{T} 的第一行，方程 $\boldsymbol{Y}=P^{\mathrm{T}}\boldsymbol{X}$ 表明：

$$y_1=\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{X}=c_1x_1+c_2x_2+\cdots+c_px_p$$

于是，y_1 是原始变量 $x_1,\cdots,x_p$ 的一个线性组合，其权重为本征向量 $\boldsymbol{u}_1$ 中的元素. 类似地，$\boldsymbol{u}_2$ 确定变量 y_2，依此类推.

【例题4】　铁路峡谷多谱段图像的初始数据包含 $\mathbf{R}^3$ 中的4百万个向量，其协方差矩阵是：[1]

---

1. 例题4以及习题5和6中的数据由马里兰州洛克维尔市的地球卫星公司提供.

$$S=\begin{bmatrix}2382.78 & 2611.84 & 2136.20\\2611.84 & 3106.47 & 2553.90\\2136.20 & 2553.90 & 2650.71\end{bmatrix}$$

求数据的主成分，并且列出第一主成分所确定的新变量.

解：S 的本征值及相应主成分(单位本征向量)为：

$$\lambda_1=7614.23 \qquad \lambda_2=427.63 \qquad \lambda_3=98.10$$

486

$$\boldsymbol{u}_1=\begin{bmatrix}0.5417\\0.6295\\0.5570\end{bmatrix} \qquad \boldsymbol{u}_2=\begin{bmatrix}-0.4894\\-0.3026\\0.8179\end{bmatrix} \qquad \boldsymbol{u}_3=\begin{bmatrix}0.6834\\-0.7157\\0.1441\end{bmatrix}$$

简单起见，保留两位小数，第一主成分所确定的变量是：

$$y_1=0.54x_1+0.63x_2+0.56x_3$$

这个方程用于生成本章实例介绍中的照片(d). 变量 x_1, x_2, x_3 是三个谱段下的信号强度. x_1 的值被转化成黑白灰度级，并生成照片(a). 类似地，x_2 和 x_3 的值分别生成照片(b)和(c). 在照片(d)的每个像素中，灰度级值由 y_1 计算出来，它是 x_1, x_2, x_3 的一个带权重的线性组合. 从这种意义上来看，照片(d)“呈现”了数据的第一主成分. ■

例题4 中，使用变量 y_1, y_2, y_3 对数据进行变换后，得到的协方差矩阵

$$\boldsymbol{D}=\begin{bmatrix}7614.23 & 0 & 0\\0 & 427.63 & 0\\0 & 0 & 98.10\end{bmatrix}$$

尽管 $\boldsymbol{D}$ 显然比原来的协方差矩阵 $\boldsymbol{S}$ 简单，但构造新变量的意义仍不明显. 不过，变量 y_1, y_2, y_3 的方差位于 $\boldsymbol{D}$ 的对角线上，且 $\boldsymbol{D}$ 的第一个方差显然远大于另两个. 基于这一事实，我们可以认为这些数据本质上是1 维的，而不是3 维的.

7.5.3 多元数据的降维

主成分分析在某些应用中非常有价值. 这类应用当中，数据的大部分方差，或者说浮动范围，都归因于极少数新变量 $y_1, \cdots, y_p$ 的方差.

可以证明，正交变量替换 $\boldsymbol{X}=\boldsymbol{PY}$ 不改变数据的全变差(严格地说，这是因为左乘以 $\boldsymbol{P}$ 不会改变向量的长度以及向量之间的夹角. 见习题12). 这就意味着，如果 $\boldsymbol{S}=\boldsymbol{PDP}^{\mathrm{T}}$，则

$$\{x_1,\cdots,x_p\text{的全变差}\}=\{y_1,\cdots,y_p\text{的全变差}\}=\mathrm{tr}(\boldsymbol{D})=\lambda_1+\cdots+\lambda_p$$

y_j 的方差是 λ_j，并且商 $\lambda_j/\mathrm{tr}(\boldsymbol{S})$ 度量了全变差中被 y_j “解释”或“捕获”的那部分.

【例题5】 计算铁路峡谷数据呈现在主成分照片(d)~(f)中的各种方差的比例，照片见本章实例介绍.

487

解：数据的全变差是

$$\mathrm{tr}(\boldsymbol{D})=7614.23+427.63+98.10=8139.96$$

[验证这个数也等于 $\mathrm{tr}(\boldsymbol{S})$.]主成分所解释的全变差比例为：

第一成分	第二成分	第三成分
$\frac{7614.23}{8139.96}=93.5\%$	$\frac{427.63}{8139.96}=5.3\%$	$\frac{98.10}{8139.96}=1.2\%$

在某种意义上，探测卫星在铁路峡谷区域所采集信息的 93.5% 呈现在照片(d)中，另有 5.3% 呈现在(e)中，仅剩下 1.2% 呈现在(f)中. ■

例题 5 的计算表明，数据在第三个(新)坐标下几乎没有方差. y_3的值非常接近零. 从几何意义来看，这些数据点分布在平面 $y_3=0$ 的附近，只要知道 y_1和 y_2的值，就可以相当精确地确定它们的位置. 事实上，y_2也有相对小的方差，这意味着数据点近似地沿直线分布，基本上是一维的. 如图 7-19，其中数据的分布看起来像根冰棍.

7.5.4　主成分变量的特征

如果 $y_1,\cdots,y_p$由 $p\times N$ 观测矩阵的主成分分析得到，则从下面意义来说，y_1有最大的方差：如果 $\boldsymbol{u}$ 是任意单位向量，且 $y=\boldsymbol{u}^{\mathrm{T}}\boldsymbol{X}$，则当 $\boldsymbol{X}$ 取遍原始数据 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_N$时，$y$ 值的方差是 $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{u}$. 根据 7.3 节定理 8，$\boldsymbol{u}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{u}$ 在全体单位向量 $\boldsymbol{u}$ 上的最大值是 $\boldsymbol{S}$ 的最大本征值 λ_1，并且这个方差在 $\boldsymbol{u}$ 等于对应本征向量 $\boldsymbol{u}_1$时取到. 类似地，定理 8 表明，在全体与 y_1 不相关的变量 $y=\boldsymbol{u}^{\mathrm{T}}\boldsymbol{X}$ 中，y_2有最大方差. 类似还有，在全体与 y_1 和 y_2 都不相关的变量中，y_3有最大方差，依此类推.

注记

实际应用中，奇异值分解是进行主成分分析的主要工具. 如果 $\boldsymbol{B}$ 是具有平均偏差形式的 $p\times N$ 观测矩阵，且 $\boldsymbol{A}=(1/\sqrt{N-1})\boldsymbol{B}^{\mathrm{T}}$，则 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 就是协方差矩阵 $\boldsymbol{S}$. $\boldsymbol{A}$ 的奇异值的平方是 $\boldsymbol{S}$ 的 p 个本征值，$\boldsymbol{A}$ 的右奇异向量是数据的主成分.

如 7.4 节所言，$\boldsymbol{A}$ 的 SVD 迭代算法比 $\boldsymbol{S}$ 的本征值分解更快而且更精确. 在本章实例介绍中提到的超光谱图像处理($p=224$)中，这一点尤其明显. 在专门的工作站上进行主成分分析，只需几秒即可完成.

488

更多阅读

Lillesand, Thomas M., and Ralph W. Kiefer, *Remote Sensing and Image Interpretation*, 4th ed. (New York: John Wiley, 2000).

基础练习

下表列出了五个男孩的体重和身高：

男孩	#1	#2	#3	#4	#5
体重(磅)	120	125	125	135	145
身高(英寸)	61	60	64	68	72

1. 求数据的协方差矩阵.
2. 对数据进行主成分分析，求解释数据中方差最多的单个数量指标[1].

习题 7.5

将习题 1 和 2 中的观测矩阵转换成平均偏差形式，并构造其样本协方差矩阵.

1. $\begin{bmatrix} 19 & 22 & 6 & 3 & 2 & 20 \\ 12 & 6 & 9 & 15 & 13 & 5 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 5 & 2 & 6 & 7 & 3 \\ 3 & 11 & 6 & 8 & 15 & 11 \end{bmatrix}$

1. 即第一主成分所确定的新变量. ——译者注

3. 求习题 1 中数据的主成分.
4. 求习题 2 中数据的主成分.
5. [**M**] 位于佛罗里达州的 Homestead 空军基地(在 1992 年遭安德鲁飓风袭击之后)制作了一幅包含三个谱成分的探测卫星图像. 数据的协方差矩阵如下所示. 求数据的第一主成分，并计算这一成分所包含全变差的比例.

$$\boldsymbol{S}=\begin{bmatrix}164.12 & 32.73 & 81.04\\ 32.73 & 539.44 & 249.13\\ 81.04 & 249.13 & 189.11\end{bmatrix}$$

6. [**M**]下面的协方差矩阵由华盛顿州哥伦比亚河的一幅探测卫星图像得到，采用了三个谱段下的数据. 以 x_1,x_2,x_3记图像中每个像素的光谱成分. 求一个形如 $y_1=c_1x_1+c_2x_2+c_3x_3$的新变量，使之在约束条件 $c_1^2+c_2^2+c_3^2=1$ 下有最大方差. 由 y_1解释的数据全变差的比例是多少?

$$\boldsymbol{S}=\begin{bmatrix}29.64 & 18.38 & 5.00\\ 18.38 & 20.82 & 14.06\\ 5.00 & 14.06 & 29.21\end{bmatrix}$$

7. 以 x_1,x_2记习题 1 中 2 维数据变量. 求一个形如 $y_1=c_1x_1+c_2x_2$，其中 $c_1^2+c_2^2=1$，使 y_1在给定的数据上有最大方差. y_1解释了数据中的多少方差?
8. 对习题 2 中的数据，重复习题 7.
9. 假设在一个大学生随机样本上进行三个测试. 设 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_N$是 $\mathbf{R}^3$ 中的观测向量，分别记录了每个学生的三次成绩，并且对 $j=1,2,3$，以 x_j记学生在第 j 次考试的成绩. 设数据的协方差矩阵是

$$\boldsymbol{S}=\begin{bmatrix}5 & 2 & 0\\ 2 & 6 & 2\\ 0 & 2 & 7\end{bmatrix}$$

设 y 是衡量学生表现的“指标”，$y=c_1x_1+c_2x_2+c_3x_3$且 $c_1^2+c_2^2+c_3^2=1$. 试选取 c_1,c_2,c_3，使得 y 在数据集上的方差最大. [提示：样本协方差矩阵的本征值是 $\lambda=3,6$ 和 9.]

489

10. [**M**]对 $\boldsymbol{S}=\begin{bmatrix}5 & 4 & 2\\ 4 & 11 & 4\\ 2 & 4 & 5\end{bmatrix}$重复习题 9.
11. 已知($\mathbf{R}^p$中)多元数据 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_N$具有平均偏差形式，设 $\boldsymbol{P}$ 是一个 $p\times p$ 矩阵，对 $k=1,\cdots,N$，定义 $\boldsymbol{Y}_k=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{X}_k$.
    a. 证明 $\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_N$具有平均偏差形式. [提示：设 $\boldsymbol{w}$ 是 $\mathbf{R}^N$中各元素均为 1 的向量. 则 $[\boldsymbol{X}_1\cdots\boldsymbol{X}_N]\boldsymbol{w}=\boldsymbol{0}$($\mathbf{R}^p$中的零向量).]
    b. 证明：如果 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_N$的协方差矩阵是 $\boldsymbol{S}$，则 $\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_N$的协方差矩阵是 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{SP}$.
12. 设 $\boldsymbol{X}$ 是一个向量，取值范围为某个 $p\times N$ 观测矩阵中的列. $\boldsymbol{P}$ 是一个 $p\times p$ 正交矩阵. 证明：变量替换 $\boldsymbol{X}=\boldsymbol{PY}$ 不改变数据的全变差. [提示：根据习题 11，只需证明 $\mathrm{tr}(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{SP})=\mathrm{tr}(\boldsymbol{S})$. 可以利用 5.4 节习题 25 提到的迹的性质.]
13. 样本协方差矩阵是对含 N 个观测数量，比如说 $t_1,\cdots,t_N$，这样一个样本方差公式的推广. 如果 m 是 $t_1,\cdots,t_N$的平均值，则样本方差由下式给出：

$$\frac{1}{N-1}\sum_{k=1}^{n}(t_k-m)^2 \tag{1}$$

试说明，例题 3 前面定义的样本协方差矩阵 $\boldsymbol{S}$ 怎样写成类似(1)的形式. [提示：利用分

块矩阵乘法将 $\boldsymbol{S}$ 写成 $1/(N-1)$ 乘以 N 个 $p\times p$ 矩阵的和. 对 $1\leqslant k\leqslant N$, 将 $\hat{\boldsymbol{X}}_k$ 写作 $\boldsymbol{X}_k-\boldsymbol{M}$.]

基础练习答案

1. 首先将数据整理成平均偏差形式. 容易看出，样本平均向量是 $\boldsymbol{M}=\begin{bmatrix}130\\65\end{bmatrix}$. 从观测向量中减去 $\boldsymbol{M}$ 得

$$\boldsymbol{B}=\begin{bmatrix}-10 & -5 & -5 & 5 & 15\\ -4 & -5 & -1 & 3 & 7\end{bmatrix}$$

于是，样本协方差矩阵是

$$\boldsymbol{S}=\frac{1}{5-1}\begin{bmatrix}-10 & -5 & -5 & 5 & 15\\ -4 & -5 & -1 & 3 & 7\end{bmatrix}\begin{bmatrix}-10 & -4\\ -5 & -5\\ -5 & -1\\ 5 & 3\\ 15 & 7\end{bmatrix}=\frac{1}{4}\begin{bmatrix}400 & 190\\ 190 & 100\end{bmatrix}=\begin{bmatrix}100.0 & 47.5\\ 47.5 & 25.0\end{bmatrix}$$

2. $\boldsymbol{S}$ 的本征值(保留两位小数)是

$$\lambda_1=123.02 \text{ 和 } \lambda_2=1.98$$

λ_1对应的单位本征向量是 $\boldsymbol{u}=\begin{bmatrix}0.900\\0.436\end{bmatrix}$. (由于 $\boldsymbol{S}$ 是 2×2 矩阵，如果没有矩阵软件，我们可以通过手算完成这一计算.)所求的数量指标是

$$y=0.900\hat{w}+0.436\hat{h}$$

其中 $\hat{w}$ 和 $\hat{h}$ 分别是平均偏差形式的体重和身高. 该指标在数据集上的方差是 123.02. 由于全变差是 $\text{tr}(\boldsymbol{S})=100+25=125$，所以该指标几乎解释了数据的全部(98.4%)方差.

490

图 7-21 绘出了基础练习 1 的原始数据以及由第一主成分 $\boldsymbol{u}$ 所确定的直线(直线用参数向量的形式表示为 $\boldsymbol{x}=\boldsymbol{M}+t\boldsymbol{u}$). 可以证明，这条直线是数据的最佳逼近，使得数据点到直线垂直距离的平方和达到最小. 事实上，主成分分析与所谓的正交回归是等价的，不过那是另一回事，以后我们可能会遇到.

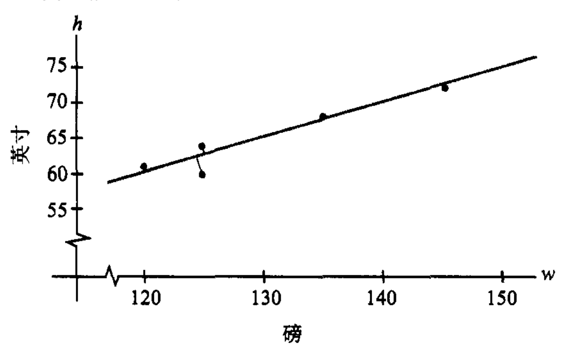


图 7-21 由数据第一主成分所确定的正交回归直线

第 7 章补充题

1. 判断下列各命题的真假，并说明理由其中 A 表示一个 $n\times n$ 矩阵.
   a. 若 A 可正交对角化，则 A 是对称的.
   b. 若 A 是正交矩阵，则 A 是对称的.
   c. 若 A 是正交矩阵，则对任意 $x\in\mathbf{R}^n$ 都有 $\|Ax\|=\|x\|$.
   d. 二次型 x^TAx 的主轴可以是 P 的列，其中 P 是任意对角化 A 的矩阵.
   e. 若 P 是各列正交的 $n\times n$ 矩阵，则 $P^T=P^{-1}$.
   f. 若某二次型的所有系数都为正数，则该二次型正定.
   g. 若对某个 x 有 $x^TAx>0$，则二次型 x^TAx 正定.
   h. 任何二次型通过一个适当的变量替换，都可以变成一个不含交叉乘积项的二次型.
   i. 二次型 x^TAx 在约束条件 $\|x\|=1$ 下的最大值是 A 对角线上的最大元.
   j. 一个正定二次型 x^TAx 的最大值是 A 的最大本征值.
   k. 正定二次型通过一个适当的变量替换 $x=Pu$，可以变成负定二次型，其中 P 是正交矩阵.
   l. 不定二次型的本征值是不确定的.
   m. 若 P 是 $n\times n$ 正交矩阵，则变量替换 $x=Pu$ 将 x^TAx 变成对应矩阵为 $P^{-1}AP$ 的二次型.
   n. 若 U 是各列正交的 $m\times n$ 矩阵，则 UU^Tx 是 x 在 Col U 上的正交投影.
   o. 如果 B 是 $m\times n$ 矩阵，x 是 $\mathbf{R}^n$ 中的单位向量，则 $\|Bx\|\leqslant\sigma_1$，其中 σ_1 是 B 的第一个奇异值.
   p. $m\times n$ 矩阵 B 的一个奇异值分解可以写作 $B=P\Sigma Q$，其中 P 是 $m\times m$ 正交矩阵，Q 是 $n\times n$ 正交矩阵，Σ 是 $m\times n$ “对角”矩阵.
   q. 若 A 是 $n\times n$ 矩阵，则 A 与 A^TA 有相同的奇异值.
2. 设 $\{u_1,\cdots,u_n\}$ 是 $\mathbf{R}^n$ 的一组标准正交基，$\lambda_1,\cdots,\lambda_n$ 是任意实数. 定义

$$A=\lambda_1u_1u_1^T+\cdots+\lambda_nu_nu_n^T$$

491

   a. 证明 A 是对称的.
   b. 证明 $\lambda_1,\cdots,\lambda_n$ 是 A 的本征值.
3. 设 A 是秩为 r 的 $n\times n$ 对称矩阵. 说明为什么 A 的谱分解将 A 表示成 r 个秩为 1 的矩阵之和.
4. 设 A 是 $n\times n$ 对称矩阵.
   a. 证明 $(\text{Col }A)^\perp=\text{Nul }A$. [提示：参考 6.1 节.]
   b. 证明：任意 $y\in\mathbf{R}^n$ 都可以写成 $y=\hat{y}+z$ 的形式，其中 $\hat{y}\in\text{Col }A$ 且 $z\in\text{Nul }A$.
5. 证明：如果 v 是 $n\times n$ 矩阵 A 的一个本征向量，且 v 对应于 A 的一个非零本征值，则 $v\in\text{Col }A$. [提示：利用本征向量的定义.]
6. 设 A 是 $n\times n$ 对称矩阵. 利用习题 5 以及 $\mathbf{R}^n$ 的一组本征向量基，给出习题 4(b) 中分解式的另一种证明.
7. 证明：$m\times n$ 矩阵 A 正定当且仅当 A 有楚列斯基分解，即，存在对角线元素为正的可逆上三角矩阵 R，使 $A=R^TR$. [提示：利用 QR 分解以及 7.2 节的习题 26.]
8. 利用习题 7 证明，若 A 正定，则 A 有 LU 分解 $A=LU$，其中 U 的对角线上有正主元.（该命题反过来也正确.）

   若 A 是矩阵，则矩阵 $G=A^TA$ 称作 A 的格拉姆矩阵. 此时，G 的元素是 A 中列的内积.
9. 证明：任意矩阵 A 的格拉姆矩阵都是半正定的，并且它的秩等于 A 的秩（见 6.5 节习题）.
10. 证明：如果 $n\times n$ 矩阵 G 是半正定的，并且秩为 r，则它是某个 $r\times n$ 矩阵 A 的格拉姆矩阵. 我们称之为 G 的显秩分解. [提示：考虑 G 的谱分解，并且先将 G 写成 BB^T，其中 B 是

$n\times r$矩阵.]

11. 证明：任意 $n\times n$ 矩阵 $\boldsymbol{A}$ 都有形如 $\boldsymbol{A}=\boldsymbol{PQ}$ 的极分解，其中 $\boldsymbol{P}$ 是与 $\boldsymbol{A}$ 等秩的 $n\times n$ 半正定矩阵，$\boldsymbol{Q}$ 是 $n\times n$ 正交矩阵. [提示：利用奇异值分解 $\boldsymbol{A}=\boldsymbol{U}\Sigma\boldsymbol{V}^{\mathrm{T}}$，并且注意 $\boldsymbol{A}=(\boldsymbol{U}\Sigma\boldsymbol{U}^{\mathrm{T}})(\boldsymbol{U}\boldsymbol{V}^{\mathrm{T}})$.] 机械工程学中，这类分解可用于模拟材料的变形. 矩阵 $\boldsymbol{P}$ 刻画了材料(在 $\boldsymbol{P}$ 的本征向量方向上)的伸长或收缩，$\boldsymbol{Q}$ 刻画了材料在空间中的旋转.

习题12～14所讨论的 $m\times n$ 矩阵 $\boldsymbol{A}$ 有简化奇异值分解 $\boldsymbol{A}=\boldsymbol{U}_r\boldsymbol{D}\boldsymbol{V}_r^{\mathrm{T}}$，以及伪逆 $\boldsymbol{A}^+=\boldsymbol{V}_r\boldsymbol{D}^{-1}\boldsymbol{U}_r^{\mathrm{T}}$.

12. 验证 $\boldsymbol{A}^+$ 的下列性质：
    a. 对任意 $\boldsymbol{y}\in\mathbb{R}^m$，$\boldsymbol{AA}^+\boldsymbol{y}$ 都是 $\boldsymbol{y}$ 在 Col $\boldsymbol{A}$ 上的正交投影.
    b. 对任意 $\boldsymbol{x}\in\mathbb{R}^n$，$\boldsymbol{A}^+\boldsymbol{Ax}$ 都是 $\boldsymbol{x}$ 在 Row $\boldsymbol{A}$ 上的正交投影.
    c. $\boldsymbol{AA}^+\boldsymbol{A}=\boldsymbol{A}$ 且 $\boldsymbol{A}^+\boldsymbol{AA}^+=\boldsymbol{A}^+$.
13. 假设方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 是相容的，令 $\boldsymbol{x}^+=\boldsymbol{A}^+\boldsymbol{b}$. 根据6.3节习题23，Row $\boldsymbol{A}$ 中存在唯一的向量 $\boldsymbol{p}$ 使得 $\boldsymbol{Ap}=\boldsymbol{b}$. 下列步骤证明了 $\boldsymbol{x}^+=\boldsymbol{p}$，并且 $\boldsymbol{x}^+$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小长度解.
    a. 证明 $\boldsymbol{x}^+\in$ Row $\boldsymbol{A}$. [提示：选取某个 $\boldsymbol{x}$，将 $\boldsymbol{b}$ 写作 $\boldsymbol{Ax}$，然后利用习题12.]
    b. 证明 $\boldsymbol{x}^+$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的一个解.
    c. 证明：若 $\boldsymbol{u}$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的任意解，则 $\|\boldsymbol{x}^+\|\leqslant\|\boldsymbol{u}\|$，并且仅当 $\boldsymbol{u}=\boldsymbol{x}^+$ 时等号成立.
14. 任给 $\boldsymbol{b}\in\mathbb{R}^m$，改造习题13以证明：$\boldsymbol{A}^+\boldsymbol{b}$ 是有最小长度的最小二乘解. [提示：考虑方程 $\boldsymbol{Ax}=\hat{\boldsymbol{b}}$，其中 $\hat{\boldsymbol{b}}$ 是 $\boldsymbol{b}$ 在 Col $\boldsymbol{A}$ 上的正交投影.]

[M]构造习题15和16中 $\boldsymbol{A}$ 的伪逆. 可以先用矩阵软件生成 $\boldsymbol{A}$ 的 SVD，如果得不到，就正交对角化 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$. 利用所得的伪逆求解 $\boldsymbol{Ax}=\boldsymbol{b}$，其中 $\boldsymbol{b}=(6,-1,-4,6)$，以 $\hat{\boldsymbol{x}}$ 记方程的解. 通过计算，验证 $\hat{\boldsymbol{x}}\in$ Row $\boldsymbol{A}$. 选取 Nul $\boldsymbol{A}$ 中的非零向量 $\boldsymbol{u}$，验证 $\|\hat{\boldsymbol{x}}\|<\|\hat{\boldsymbol{x}}+\boldsymbol{u}\|$，由习题13(c)知这个不等式是正确的.

15. $\boldsymbol{A}=\begin{bmatrix}-3 & -3 & -6 & 6 & 1\\ -1 & -1 & -1 & 1 & -2\\ 0 & 0 & -1 & 1 & -1\\ 0 & 0 & -1 & 1 & -1\end{bmatrix}$

16. $\boldsymbol{A}=\begin{bmatrix}4 & 0 & -1 & -2 & 0\\ -5 & 0 & 3 & 5 & 0\\ 2 & 0 & -1 & -2 & 0\\ 6 & 0 & -3 & -6 & 0\end{bmatrix}$

492

附录 A　简化阶梯形式的唯一性

【定理】　简化阶梯形式的唯一性

每一个矩阵 $\boldsymbol{A}$ 都行等价于唯一一个简化阶梯矩阵 $\boldsymbol{U}$.

证明：证明要用到 4.3 节中关于行等价矩阵的列有相同线性相关关系的思想.

行化简算法表明，至少存在一个这样的矩阵 $\boldsymbol{U}$. 假设 $\boldsymbol{A}$ 与简化阶梯矩阵 $\boldsymbol{U}$ 和 $\boldsymbol{V}$ 行等价. $\boldsymbol{U}$ 的每一行中最左边的非零元是"首项 1". 把首项 1 所在的位置称为主元位置，称含有主元位置的列为主元列.（该定义仅用了 $\boldsymbol{U}$ 和 $\boldsymbol{V}$ 的阶梯性质，并没有假定简化阶梯形式的唯一性.）

$\boldsymbol{U}$ 和 $\boldsymbol{V}$ 的主元列正好是那些与其左边各列线性无关的非零列（第一列如果是非零列，则自动满足这种情况）. 由于 $\boldsymbol{U}$ 和 $\boldsymbol{V}$ 行等价（都由 $\boldsymbol{A}$ 通过行变换得到），它们的列有相同的线性相关关系. 因此，$\boldsymbol{U}$ 和 $\boldsymbol{V}$ 的主元列出现在相同的位置上. 且由于它们是简化阶梯形式，如果有 r 个这样的列，主元列就是 $m \times m$ 单位矩阵中的前 r 列. 因此，$\boldsymbol{U}$ 和 $\boldsymbol{V}$ 对应的主元列相等.

最后，考虑 $\boldsymbol{U}$ 中任意一个非主元列，比方说第 j 列. 该列或者为零，或者是其左边各主元列的线性组合（因为那些主元列是第 j 列左边各列所张成的子空间的一组基）. 不论是哪种情况，都可以写成 $\boldsymbol{Ux}=\boldsymbol{0}$，其中 $\boldsymbol{x}$ 的第 j 个元素为 1. 同样也有 $\boldsymbol{Vx}=\boldsymbol{0}$，表示 $\boldsymbol{V}$ 的第 j 列或者为零，或者是其左边各列的同一线性组合. 由于 $\boldsymbol{U}$ 和 $\boldsymbol{V}$ 对应的主元列相等，$\boldsymbol{U}$ 和 $\boldsymbol{V}$ 的第 j 列也相等. 对任意非主元列上述结论都成立，因此 $\boldsymbol{U}=\boldsymbol{V}$，这就表明 $\boldsymbol{U}$ 是唯一的. ■

附录B 复　　数

复数(complex number)是写成如下形式的一个数

$$z = a + bi$$

其中 a 和 b 为实数，i 为满足关系 $i^2 = -1$ 的一个形式符号. 称 a 为 z 的**实部**(real part)，记作 Re z;b 为 z 的**虚部**(imaginary part)，记作 Im z. 两个复数相等当且仅当它们的实部和虚部分别相等. 例如，如果 $z = 5 + (-2)i$，则 Re $z = 5$，Im $z = -2$. 简单起见，我们记为 $z = 5 - 2i$.

实数 a 可以看成一类特殊的复数，即 $a + 0i$. 此外，实数的算术运算也可以推广到复数集合上.

复数系(complex number system)，记作 **C**，是由全体复数构成的集合，并且定义了下列加法和乘法运算：

$$\boxed{(a + bi) + (c + di) = (a + c) + (b + d)i} \tag{1}$$

$$\boxed{(a + bi)(c + di) = (ac - bd) + (ad + bc)i} \tag{2}$$

当(1)和(2)中的 b 和 d 取零时，上述法则就是实数的加法和乘法法则. 容易验证，**R** 中常用法则在 **C** 中也成立. 因此，复数的乘法通常利用代数展开来计算，如下面例题所示.

【例题 1】 $(5 - 2i)(3 + 4i) = 15 + 20i - 6i - 8i^2 = 15 + 14i - 8(-1) = 23 + 14i$

即，将 $5 - 2i$ 的每一项乘以 $3 + 4i$ 的每一项，然后利用 $i^2 = -1$，把结果整理成 $a + bi$ 的形式. ■

复数 z_1 和 z_2 之间的减法定义为

$$z_1 - z_2 = z_1 + (-1)z_2$$

特别地，我们用 $-z$ 代替 $(-1)z$.

$z = a + bi$ 的**共轭**(conjugate)是一个复数 $\bar{z}$(读作“z - bar”)，定义为

$$\boxed{\bar{z} = a - bi}$$

将 z 的虚部符号取反，就得到了 $\bar{z}$.

【例题 2】 $-3 + 4i$ 的共轭是 $-3 - 4i$；记为 $\overline{-3 + 4i} = -3 - 4i$. ■

注意，如果 $z = a + bi$，则

$$z\bar{z} = (a + bi)(a - bi) = a^2 - abi + bai - b^2i^2 = a^2 + b^2 \tag{3}$$

$z\bar{z}$为非负实数，因而有平方根. z 的**绝对值**(absolute value)或**模**(modulus)为实数 $|z|$，定义如下

$$\boxed{|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}}$$

如果 z 为实数，则 $z = a + 0i$，且 $|z| = \sqrt{a^2}$，就是通常意义下 a 的绝对值.

下面列出了一些有关共轭和绝对值的常用性质，其中 w 和 z 表示复数.

1. $\bar{z} = z$ 当且仅当 z 为实数.
2. $\overline{w + z} = \bar{w} + \bar{z}$.

3. $\overline{wz}=\overline{w}\,\overline{z}$；特别地，如果 r 为实数，则$\overline{rz}=r\,\overline{z}$.

4. $z\overline{z}=|z|^2\geqslant 0$.

5. $|wz|=|w||z|$.

6. $|w+z|\leqslant|w|+|z|$.

如果 $z\neq 0$，则$|z|>0$ 且 z 有乘法逆元，记为 $1/z$ 或 z^{-1}，由下式定义：

$$\boxed{\frac{1}{z}=z^{-1}=\frac{\overline{z}}{|z|^2}}$$

商 w/z 是 $w(1/z)$ 的简写形式.

【例题 3】 设 $w=-3+4\mathrm{i}$，$z=5-2\mathrm{i}$. 计算 $z\overline{z}$，$|z|$和 w/z.

解：根据(3)，

$$z\overline{z}=5^2+(-2)^2=25+4=29$$

所求的绝对值为$|z|=\sqrt{z\overline{z}}=\sqrt{29}$. 为了计算 w/z，首先在分子和分母上同乘以分母的共轭 $\overline{z}$. 根据(3)，这个步骤可以消去分母中的 i：

$$\frac{w}{z}=\frac{3+4\mathrm{i}}{5-2\mathrm{i}}=\frac{3+4\mathrm{i}}{5-2\mathrm{i}}\cdot\frac{5+2\mathrm{i}}{5+2\mathrm{i}}=\frac{15+6\mathrm{i}+20\mathrm{i}-8}{5^2+(-2)^2}=\frac{7+26\mathrm{i}}{29}=\frac{7}{29}+\frac{26}{29}\mathrm{i}$$ ■

几何解释

每个复数 $z=a+b\mathrm{i}$ 都与平面 $\mathbb{R}^2$ 中的一点(a,b)对应，如图 B-1 所示. 因为点$(a,0)$对应于实数，所以水平轴被称为**实轴**(real axis). 点$(0,b)$对应于形如 $0+b\mathrm{i}$ 的**纯虚数**(pure imaginary numbers)，或简写为 $b\mathrm{i}$，所以竖轴被称为**虚轴**(imaginary axis). z 的共轭是 z 关于实轴的镜像. z 的绝对值是从(a,b)到原点的距离.

复数 $z=a+b\mathrm{i}$ 和 $w=c+d\mathrm{i}$ 的加法对应于 $\mathbb{R}^2$ 中向量(a,b)和(c,d)的加法，如图 B-2 中所示.

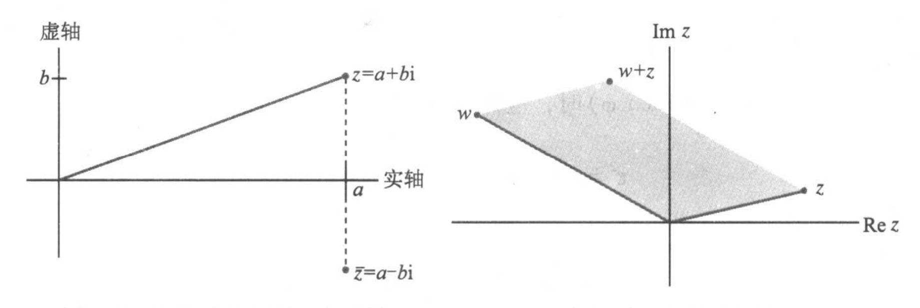


图 B-1 复数的共轭是一个镜像　　图 B-2 复数的加法

为了给出复数乘法的几何表示，我们要用到 $\mathbb{R}^2$ 的极坐标. 给定非零复数 $z=a+b\mathrm{i}$，设 φ 为实轴正向与点(a,b)间的夹角，如图 B-3 所示，$-\pi<\varphi\leqslant\pi$. 角度 φ 称为 z 的**幅角**(argument)，记为 $\varphi=\arg z$. 由三角函数知识有，

$$a=|z|\cos\varphi,\quad b=|z|\sin\varphi$$

所以

$$z=a+b\mathrm{i}=|z|(\cos\varphi+\mathrm{i}\sin\varphi)$$

如果 w 为另一个非零复数，比如说

$$w=|w|(\cos\vartheta+\mathrm{i}\sin\vartheta)$$

然后应用正弦、余弦两角和的标准三角恒等式，我们可以验证

$$wz = |w||z|[\cos(\vartheta+\varphi) + \mathrm{i}\sin(\vartheta+\varphi)] \tag{4}$$

见图 B-4. 对商也可以写出一个类似的极坐标公式. 乘积和商的公式可以用文字描述如下：

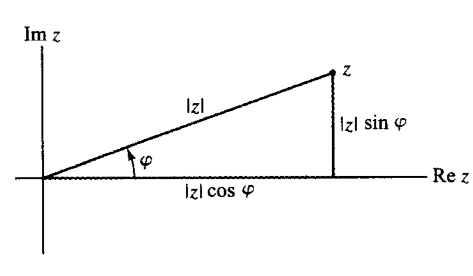


图 B-3　z 的极坐标　　　　图 B-4　极坐标的乘法

> 在极坐标的形式下，两个非零复数乘积的绝对值和辐角分别等于它们的绝对值之积、辐角之和；两个非零复数的商的绝对值和辐角分别等于它们的绝对值的商、辐角之差.

【例题 4】

a. 如果 w 的绝对值为 1，则 $w = \cos\vartheta + \mathrm{i}\sin\vartheta$，其中 ϑ 为 w 的幅角. 任意非零复数 z 乘以 w 就相当于将 z 旋转角度 ϑ.

b. i 自身的幅角为$\pi/2$，因此 z 乘以 i 就是将 z 旋转了$\pi/2$ 角度（见图 B-5）. 例如，将 $3+\mathrm{i}$ 旋转到$(3+\mathrm{i})\mathrm{i} = -1+3\mathrm{i}$. ■

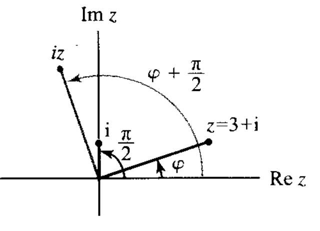


图 B-5　乘以 i

复数的幂

当 $z = w = r(\cos\varphi + \mathrm{i}\sin\varphi)$时，应用公式(4)，有

$$z^2 = r^2(\cos 2\varphi + \mathrm{i}\sin 2\varphi)$$

且

$$\begin{aligned} z^3 &= z \cdot z^2 \\ &= r(\cos\varphi + \mathrm{i}\sin\varphi) \cdot r^2(\cos 2\varphi + \mathrm{i}\sin 2\varphi) \\ &= r^3(\cos 3\varphi + \mathrm{i}\sin 3\varphi) \end{aligned}$$

一般来说，对任意的正整数 k，

$$z^k = r^k(\cos k\varphi + \mathrm{i}\sin k\varphi)$$

这就是棣莫弗定理.

复数与 R^2

尽管 $\mathbf{R}^2$（见图 B-6）与 $\mathbf{C}$（见图 B-7）中的元素是一一对应的，并且加法运算本质上一样，但 $\mathbf{R}^2$和 $\mathbf{C}$ 之间有着逻辑上的区别. 在 $\mathbf{R}^2$中，我们只可以用一个实数量乘以一个向量，然而在 $\mathbf{C}$ 中我们可以把任意两个复数相乘得到第三个复数（不

考虑 $\mathbf{R}^2$中的点积，因为它得到的是数量，并非 $\mathbf{R}^2$中的元素）．为了强调这个区别，我们对 **C** 中的元素使用数量的记法．

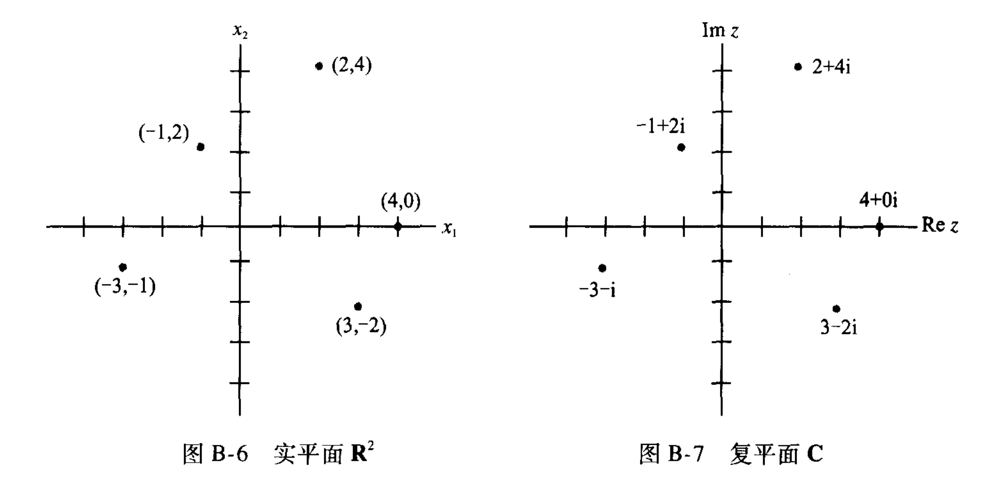


图 B-6 实平面 $\mathbf{R}^2$　　图 B-7 复平面 **C**